Remarks

Reconsideration and withdrawal of the rejections of the claims, in view of the amendments and remarks herein, is respectfully requested. The amendments are intended to advance the application and are not intended to concede to the correctness of the Examiner's position or to prejudice the prosecution of the claims prior to amendment, which claims are present in a continuation of the above-referenced application. Claim 66 is amended, and claims 90-92 are added. Claims 66-67, 70-74, 88 and 90-92 are now pending in this application.

Amended claim 66 is supported by page 4, lines 29-32 and page 25, lines 27-30 of the specification.

New claim 90 is supported by Figure 3.

New claims 91-92 are supported at page 9, lines 14-17 of the specification.

The 35 U.S.C. § 101 and § 112 "Utility" Rejections

The Examiner rejected claims 66-67, 70-74 and 88 under 35 U.S.C. § 101 and 35 U.S.C. § 112, first paragraph, as the claimed invention is allegedly not supported by either a specific, asserted utility or a well established utility. These rejections are respectfully traversed.

The Utility Examination Guidelines of the U.S. Patent and Trademark Office explicitly state that "the reasonable assignment of a new protein to [a] class of sufficiently conserved proteins [on the basis of sequence homology] would impute the same specific, substantial, and credible utility to the assigned protein." Utility Examination Guidelines, *Fed. Reg.* 66:1092-1099, at 1096, column 3 (2001).

The specification discloses that the nucleic acid molecules of the invention encode polypeptides that catalyze the synthesis of polyketides such as ones with a bryopyran ring (page 2, lines 17-19). To obtain those nucleic acid molecules, degenerate primers based on conserved regions of a gene encoding a beta-ketosynthase (KS) domain from type I polyketide synthases (PKS-I) were used to amplify *Bugula neritina* DNA (page 50, lines 4-5 and lines 25-27). Nine clones (clones KS_a-KS_i) obtained from the amplified DNA were sequenced (Table III). It is disclosed that all clones exhibited significant similarity to other PKS-I genes (page 53, lines 22-23).

The most prevalent clone, clone KS_a, was subjected to inverse PCR and the reaction yielded a DNA of 737 bp (SEQ ID NO:13 encoding SEQ ID NO:14). It is disclosed that SEQ ID NO:13 encoded a gene product that was similar to the 3' end of a KS domain and intermodular region upstream of an acyl transferase (AT) domain when aligned with the *pik*AI KS domain of *Streptomyces venezuelae* (citing to Xu et al., Proc. Natl. Acad. Sci. USA, 8:12111 (1998)) and DEBS1 from *Saccharomyces erythraea* (citing to Donadio et al., Science, 759:675 (1991)), both of which form a PKS-I (page 53, lines 16-20). It is also disclosed that SEQ ID NO:13 represents a single open reading frame (ORF) of a PKS-I (page 53, lines 20-21).

The KS_a clone was used as a probe to identify larger (cosmid) clones with related sequences present in a library with large inserts of *B. neritina* DNA (page 58, lines 1-19). Two cosmid clones, clones 3A and 6A, were identified and characterized, and probes derived from clone 6A were employed to rescreen the library (page 59, lines 4-6). Clones from this screening (clones 5A and 5B) were rescreened with the KS_a probe (page 59, lines 6-8). Figure 12 shows a map for clones 3A, 6A, 5A and 5B.

Clone 5A had PKS homology at both ends, while clone 5B had PKS homology at the 5' end (page 59). *Pst*I fragments of clones 5A and 5B were subcloned and sequenced. It is disclosed that for clone 5B, *Pst*I fragments A4 and B1, D4 and C1, C1 and E1, E1 and A3, and A3 and A7, overlap (SEQ ID NO:33 represents A4/B1 overlap; SEQ ID NO:34 represents D4/C1 overlap; SEQ ID NO:35 represents C1/E1 overlap and SEQ ID NO:36 represents E1/A3 overlap) (page 60, lines 10-14). Clone 5B *Pst*I A7 corresponds to SEQ ID NO:37 which, due to a restriction requirement, is the only sequence from clone 5B under examination in this application.

It is disclosed that clone 5B (which includes SEQ ID NO:37) corresponds to sequences in the 3' two-thirds of the bryostatin gene cluster (Figure 13) and encodes sequences related to PKS based on sequence homology (page 5, lines 24-25), and that SEQ ID NO:37 encodes sequences at the end of the PKS-I (page 60, lines 17-20).

Further evidence that the bryostatin PKS is structurally related to other PKS-I is provided in Davidson et al. (Appl. Environ. Microbiol., 67:4531 (2001), a copy is enclosed herewith). For example, the amino acid sequence encoded by *B. neritina* clone KS_a in Davidson et al., a clone prepared by PCR amplification with degenerate primers based on KS domains in PKS-I (which

clone is disclosed in Example 3 of the present specification), is aligned with the KS domains of other PKS-I and fatty acid synthases (FAS) (see Figure 2 of Davidson et al.). For instance, the amino acid sequence encoded at the 5' end of SEQ ID NO:37 has homology to a *B. subtilis* PKS in Accession No. NP389600. Other domains in a PKS-I, e.g., an AT, ACP, KR, DH, ER or TE domain (see Donadio et al., Science, 252:675 (1991)) and Accession No. Q03131, a copy of each is enclosed herewith), are well known and readily identifiable. Moreover, the Examiner is requested to note that an AT domain is not necessarily part of a PKS module (see, e.g., Cheng et al., Proc. Natl. Acad. Sci. USA, 100:3149 (2003), a copy is enclosed herewith).

Further, the utility for nucleic acid encoding PKS is well established, e.g., to recombinantly produce the product of the PKS (see, e.g., Kao et al., Science, 265:509 (1994), and U.S. Patent Nos. 5,849,541, 6,022,731, and 6,033,883; all of record) and, as acknowledged by the Examiner, in swapping experiments with other PKS gene clusters (page 10 of the Office Action).

As the specification provides a specific, asserted and well established utility for the claimed subject matter, withdrawal of the § 101/§ 112 utility rejections is respectfully requested.

The 35 U.S.C. § 112, Second Paragraph, Rejections

The Examiner rejected claims 66-67 and 70-74 under 35 U.S.C. § 112 second paragraph, as being indefinite. Specifically, the Examiner asserts that: 1) the phrase "at least one polypeptide that catalyzes at least one step in the synthesis of at least one polyketide or bryopyran ring" is unclear as to its metes and bounds, and that no polyketide synthase polypeptides are described as being encoded by SEQ ID NO:37; 2) the first occurrence of "or the complement thereof" in claim 66 is unclear; and 3) the hybridization conditions in claim 66 are unclear as they are disclosed as wash conditions in the specification. These rejections are respectfully traversed.

The amendments to claim 66 render bases 2) and 3) of the § 112(2) rejection moot.

With respect to 1), it is Applicant's position that the metes and bounds of the claims in view of the specification are clear. Figure 3 shows the domain structure encoded by three open reading frames for a bryopyran synthase, including the activities found in PKS-I. SEQ ID NO:37 (described in Figure 18B) is included in clone 5B (page 60, lines 17-18), which

corresponds to the 3' two-thirds of the bryostatin gene cluster, and by visual inspection has three open reading frames. Figure 13 shows the location of clone 5B in the bryostatin gene cluster and regions in the cluster which have homology to PKS. Thus, one of ordinary skill in the art in possession of Applicant's specification would be apprized of the metes and bounds of the phrase "at least one polypeptide that catalyzes at least one step in the synthesis of at least one bryopyran ring" and the domains encoded by SEQ ID NO:37. Nonetheless, to advance the application, claim 66 is amended.

Accordingly, withdrawal of the rejections under 35 U.S.C. § 112, second paragraph, is respectfully requested.

The 35 U.S.C. § 112, First Paragraph, Rejections

The Examiner rejected claims 66-67 and 70-74 under 35 U.S.C. § 112, first paragraph, for having an unclear function and without any clear structural limitation. This rejection, as it may be maintained with respect to the pending claims, is respectfully traversed.

As amended, claim 66 is directed to a composition comprising at least one isolated nucleic acid molecule that encodes at least a portion of polyketide synthase, or a complement of the nucleic acid molecule, wherein the portion includes a domain having at least one activity of a polyketide synthase type I, wherein the at least one nucleic acid molecule hybridizes to SEQ ID NO:37 or the complement thereof after washing in 0.015 M NaCl/0.0015 M sodium citrate, 0.1% SDS at 50°C. Therefore, the claims are in compliance with the written description requirement of § 112(1).

The Examiner further rejected claims 66-67 and 70-74 under 35 U.S.C. § 112, first paragraph, alleging that the specification does not reasonably provide enablement for nucleic acid molecules with the "low degree of sequence identity claimed and having a vague function." This rejection is respectfully traversed.

Specifically, the Examiner asserts that it would require undue experimentation by the art worker to use the claimed invention. The specification discloses that the following wash conditions are "stringent conditions": 0.015 M NaCl/0.0015 M sodium citrate, 0.1% SDS at 50°C. Clearly, it is well within the skill of the art worker to perform nucleic acid hybridizations under particular conditions to a specific sequence, i.e., SEQ ID NO:37 or its complement.

Moreover, it is clearly within the skill of the art worker to modify a particular nucleic acid sequence, i.e., modify SEQ ID NO:37 or its complement, for example, by random or directed mutagenesis and determine the T_m of that modified sequence, and/or to conduct hybridization experiments to determine whether the modified sequence hybridizes to a particular nucleic acid sequence under certain conditions.

Further, as discussed above, the function of a PKS-I is well known to the art worker, and it is well within the skill of the art worker to identify domains within a PKS-I by homology to domains in other PKS-I, as well as regions within those domains which are conserved (see Davidson et al.). Based on those comparisons, one of ordinary skill in the art can modified particular regions of a domain via directed mutagenesis and determine if the encoded mutant domain has altered activity. For instance, as conceded by the Examiner, the art worker would be able to use portions of a PKS in swapping experiments with other PKS gene clusters to produce polyketide analogs (page 10 of the Office Action). Thus, Applicant has fully enabled the claimed invention.

Accordingly, withdrawal of the § 112(1) written description and enablement rejections is respectfully requested.

Conclusion

Applicant respectfully submits that the claims are in condition for allowance and notification to that effect is earnestly requested. The Examiner is invited to telephone Applicant's attorney (612-373-6959) to facilitate prosecution of this application.

If necessary, please charge any additional fees or credit overpayment to Deposit Account No. 19-0743.

Respectfully submitted,

MARGO HAYGOOD ET AL.

By their Representatives,

SCHWEGMAN, LUNDBERG, WOESSNER & KLUTH, P.A.
P.O. Box 2938
Minneapolis, MN 55402
(612) 373-6959

Date February 23, 2004 By _____
Janet E. Embretson
Reg. No. 39,665

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail, in an envelope addressed to: Mail Stop AF, Commissioner of Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on this 23 day of February, 2004.

Candis B. Buending _____

Name Signature

Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis

Yi-Qiang Cheng[†‡], Gong-Li Tang[†‡], and Ben Shen[†§¶]

[†]Division of Pharmaceutical Sciences and [§]Department of Chemistry, University of Wisconsin, Madison, WI 53705

Edited by Christopher T. Walsh, Harvard Medical School, Boston, MA, and approved January 10, 2003 (received for review December 2, 2002)

Type I polyketide synthases (PKSs) are multifunctional enzymes that are organized into modules, each of which minimally contains a β-ketoacyl synthase, an acyltransferase (AT), and an acyl carrier protein. Here we report that the leinamycin (LNM) biosynthetic gene cluster from *Streptomyces atroolivaceus* S-140 consists of two PKS genes, *lnmI* and *lnmJ*, that encode six PKS modules, none of which contain the cognate AT domain. The only AT activity identified within the *lnm* gene cluster is a discrete AT protein encoded by *lnmG*. Inactivation of *lnmG*, *lnmI*, or *lnmJ in vivo* abolished LNM biosynthesis. Biochemical characterization of LnmG *in vitro* showed that it efficiently and specifically loaded malonyl CoA to all six PKS modules. These findings unveiled a previously unknown PKS architecture that is characterized by a discrete, iteratively acting AT protein that loads the extender units in trans to "AT-less" multifunctional type I PKS proteins for polyketide biosynthesis. This PKS structure provides opportunities for PKS engineering as exemplified by overexpressing *lnmG* to improve LNM production.

Polyketide natural products include many clinically important drugs such as erythromycin (antibacterial), epothilone (anticancer), rapamycin (immunosuppressant), and lovastatin (antihypercholesterolemic) (1). They are biosynthesized from short carboxylic acid precursors by polyketide synthases (PKSs) (2). Three types of bacterial PKSs are known to date. (*i*) Type I PKSs are multifunctional enzymes that are organized into modules, each of which harbors a set of distinct, noniteratively acting activities responsible for the catalysis of one cycle of polyketide chain elongation (3). (*ii*) Type II PKSs are multienzyme complexes that carry a single set of iteratively acting activities and minimally consist of the β-ketoacyl synthase (KS) α and β subunits and an acyl carrier protein (ACP) (4). [The KSβ subunit is also known as chain-length factor (5) or chain-initiation factor (6).] (*iii*) Type III PKSs, also known as chalcone synthase-like PKSs, are distributed predominately in plant and have been characterized from microorganisms only very recently (7, 8). They are essentially condensing enzymes that lack ACP and act directly on acyl CoA substrates.

Among type I PKSs characterized to date, each module minimally contains three domains, β-KS, acyltransferase (AT), and ACP, that select, activate, and catalyze a decarboxylative Claisen condensation between the extender unit and the growing polyketide chain, generating a β-ketoacyl-*S*-ACP intermediate. Optional domains are found between AT and ACP, which carry out the variable set of reductive modifications of the β-keto group before the next round of chain extension. The order of modules in the PKS enzymes dictates the sequence of biosynthetic events, and the variation of domains within the modules affords the structural diversity observed in the resultant polyketide products (2, 3). This one-to-one correspondence between the modularity of enzyme activities and the structures of the resultant products provides the molecular basis for combinatorial manipulation of type I PKSs for structural diversity (9–12).

Leinamycin (LNM, **1**) (Fig. 1) is a hybrid peptide–polyketide natural product that showed potent antitumor activity, most significantly against tumor cell lines that were resistant to clinically important anticancer drugs (13–16). We recently cloned and localized the *lnm* biosynthesis gene cluster to a 172-kb DNA region from *Streptomyces atroolivaceus* S-140, and DNA sequence analysis of the four overlapping cosmids (pBS3004, pBS3005, pBS3006, and pBS3007) revealed 72 ORFs (17). Sequential inactivation of ORFs from both ends of the sequenced region led to the assignment of the *lnm* gene cluster to consist of 27 ORFs, of which two (*lnmQ* and *lnmP*) encode nonribosomal peptide synthetase (NRPS), one (*lnmI*) encodes a hybrid NRPS–PKS enzyme, and two (*lnmJ* and *lnmG*) encode PKS. Surprisingly, all six PKS modules encoded by *lnmI* and *lnmJ* lack their cognate AT domain, and the only AT activity identified within the *lnm* gene cluster is a discrete AT protein encoded by *lnmG* (Fig. 1).

Because a PKS module cannot be functional unless its ACP domain is loaded with the extender unit (2, 3), we proposed that LnmG provides the missing AT activity in trans to the LnmI and LnmJ PKS enzymes. To test this hypothesis, we first inactivated *lnmI*, *lnmJ*, and *lnmG*, confirming that they are essential for **1** biosynthesis. We then overexpressed *lnmG* and the ACP domains from all six PKS modules either as an individual domain or multidomain PKS module, along with *lnmP*, which encodes a peptidyl carrier protein (PCP), as a negative control. We finally showed that LnmG efficiently and specifically loaded malonyl CoA to the ACP domains from all six PKS modules but not to LnmP PCP. Our findings unveiled a previously unknown PKS architecture that provides opportunities for PKS engineering as exemplified by overexpressing *lnmG* to improve **1** production.

Materials and Methods

Sequence and *in Vivo* Analysis of the *lnm* Gene Cluster. DNA sequencing, bioinformatic analyses of the sequenced 135,638-bp contiguous DNA, and functional assignments of the deduced gene products were summarized in GenBank (accession no. AF484556) (17). The boundaries of the *lnm* gene cluster were identified by gene replacement of *orf*(−13), *orf*(−11), *orf*(−2), *orf*(−1), and *lnmA* for the upstream boundary and of *orfZ'*, *orf*(+1), *orf*(+2), *orf*(+3), *orf*(+4), and *orf*(+6) for the downstream boundary, respectively. Inactivation of genes within the *lnm* gene cluster abolished **1** production, whereas that of genes outside the *lnm* gene cluster had no effect on **1** production.

LNM Production, Isolation, and Analysis. LNM production and isolation from both the wild-type and recombinant *S. atroolivaceus* strains were carried out as reported (17). HPLC analysis was

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: PKS, polyketide synthase; KS, β-ketoacyl synthase; ACP, acyl carrier protein; AT, acyltransferase; LNM, leinamycin; NRPS, nonribosomal peptide synthetase; PCP, peptidyl carrier protein; KR, ketoreductase; Svp, *Streptomyces verticillus* phosphopantetheinyl transferase; ESI-MS, electrospray ionization MS.

Data deposition: The nucleotide sequence reported in this paper has been deposited in the GenBank database (accession no. AF484556).

See commentary on page 3010.

[‡]Y.-Q.C. and G.-L.T. contributed equally to this work.

[¶]To whom correspondence should be addressed at: Division of Pharmaceutical Sciences, School of Pharmacy, University of Wisconsin, 777 Highland Avenue, Madison, WI 53705. E-mail: bshen@pharmacy.wisc.edu.

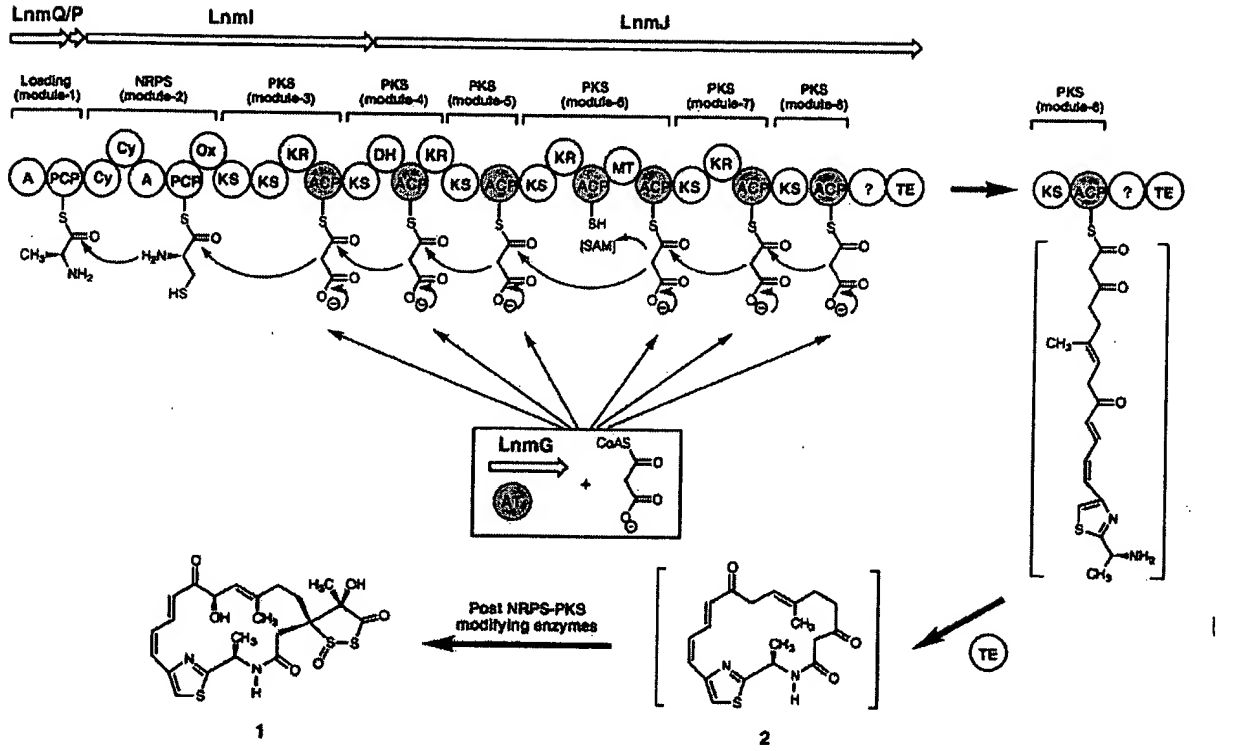

Fig. 1. Hypothesis for LNM (1) biosynthesis and modular organization of the Lnm hybrid NRPS–PKS megasynthetase with the discrete LnmG AT enzyme loading the malonyl CoA extender to all six PKS modules. The structures in brackets are hypothetical. It is not known whether one or both ACPs in module 6 are loaded with the malonyl group *in vivo*, although LnmG prefers ACP6-2 *in vitro*. Cy, condensation/cyclization; DH, dehydratase; MT, methyl transferase; Ox, oxidation; ?, domain of unknown function; TE, thioesterase.

carried out on a Microsorb-MV C-18 column (5 μm, 100 Å, 250 × 4.6 mm, Varian) eluted with a gradient from 100% buffer A (20% CH_3CN, pH 3.6 with HOAc) to 68% buffer B (80% CH_3CN, pH 3.6 with HOAc) in 40 min at a flow rate of 1 ml/min and UV detection at 320 nm.

Inactivation by Gene Replacement. To inactivate *lnmI*, *lnmJ*, and *lnmG*, internal fragments (a 5,906-bp *SphI–XhoI* fragment that harbors the condensation/cyclization, adenylation, PCP, and oxidation domains of NRPS module 2 and the first KS domain of PKS module 1 for *lnmI*, a 4,571-bp *EcoRI–NotI* fragment that harbors the ketoreductase (KR) and ACP domain of PKS module 7 and the KS domain of PKS module 8 for *lnmJ*, and a 388-bp *SalI–BamHI* fragment for *lnmG*) were replaced with the *aac(3)IV* apramycin-resistance gene (18), and the mutated *lnmI*, *lnmJ*, and *lnmG* genes were cloned into pSET151 (19) to yield pBS3017, pBS3018, and pBS3019, respectively. These constructs were introduced into *S. atroolivaceus* S-140 by conjugation (17) and selected for apramycin resistance and thiostrepton-sensitive phenotype to isolate the desired double-crossover mutant strains SB3002 (Δ*lnmI*), SB3003 (Δ*lnmJ*), and SB3004 (Δ*lnmG*), respectively. The genotypes of these mutants were confirmed by Southern analysis.

Expression of *lnm* Gene in *Escherichia coli* and Purification of the Resultant Recombinant Proteins. The *lnmG* gene, the seven ACP domains from *lnmI* and *lnmJ*, the tridomain PKS module 4 of *lnmJ*-(*DH-ACP-KR*), and the *lnmP* gene all were amplified by PCR using the following primers: for *lnmG*, 5′-CG GAA TTC CAT ATG GTG GCA CTG GTT TTC CCG-3′ and 5′-CG GCC AAG CTT GCG GCG GGC GAG GAC GTC-3′; for *lnmP*, 5′-CG GAA TTC CAT ATG TGG GAC CAC AAG TTC GAG-3′ and 5′-CG CGC AAG CTT TCG GCC GGC TCC GTC GAG-3′; for *lnm-ACP3*, 5′-CG GAA TTC CAT ATG TCA GTC ACC GGG CCG CCC-3′ and 5′-CG CGC AAG CTT CCC GAG GTC CGC CAG ATG-3′; for *lnmJ-ACP4*, 5′-CG GAA TTC CAT ATG GGG CCG GAC GCG GTG CGC-3′ and 5′-CG CGC AAG CTT GAA CTC GGC GTA CAG GTG-3′; for *lnmJ-ACP5*, 5′-CG GAA TTC *CAT* ATG GAC CCG CAG GAG GTG CTG-3′ and 5′-CG CGC AAG CTT GTG CAG TTC CCT GAC GTG-3′; for *lnmJ-ACP6-1*, 5′-CG GAA TTC CAT ATG TCG GCC GAG GCC GTG CGG-3′ and 5′-CG CGC AAG CTT GTG TTC CTG GCG GAA GTA CC-3′; for *lnmJ-ACP6-2*, 5′-CG GAA TTC CAT ATG TCG CCC GAG TCC GTG CGG-3′ and 5′-CG CGC AAG CTT GTG CTC GGC GCT CAG GTA C-3′; for *lnmJ-ACP7*, 5′-CG GAA TTC CAT ATG CTG CGG GAG CTC GTG GAG-3′ and 5′-CG CGC AAG CTT ATG GTG CTG CGT CAG GTA CT-3′; for *lnmJ-ACP8*, 5′-CG GAA TTC CAT ATG GCC GCC TCC ACC GTC GTC-3′ and 5′-CG CGC AAG CTT GAC CAG CGG CGC GAC GAA C-3′; and for *lnmJ-(DH-ACP4-KR)*, 5′-A TGA ATT CAT ATG AAC GTG CCC TCC GCA C-3′ and 5′-AT AAG CTT GCC GTC CGG GGA GTC AGG-3′. The numbers after each ACP refer to PKS modules from which they are derived with 6-1 and 6-2 to indicate the first and second ACP, respectively, for PKS module 6, and the restriction sites CAT ATG for *NdeI* and AAG CTT for *HindIII* designed in primers are underlined. The resultant products were sequenced to confirm PCR fidelity and cloned as *NdeI–HindIII* fragments into the same sites of pET28a (Novagen), yielding expression constructs pBS3020–pBS3029 for *lnmG*, *lnmP*, *lnmJ-ACP3*, *lnmJ-ACP4*, *lnmJ-ACP5*, *lnmJ-ACP6-1*, *lnmJ-ACP6-2*, *lnmJ-ACP7*, *lnmJ-ACP8*, and *lnmJ-(DH-ACP4-KR)*, respectively. Introduction of pBS3020–pBS3029 into *E. coli* BL-21 (DE-3) resulted in the overproduction of these gene products as His_6-tagged fusion proteins. The latter were purified by affinity chromatography on

Ni-nitrilotriacetic acid resin (Qiagen, Valencia, CA), dialyzed against 25 mM Tris·HCl, pH 7.0 (for LnmG) or pH 8.0 (for ACPs or LnmP), 25 mM NaCl/10% glycerol/2 mM DTT, and stored at −80°C.

In Vitro Assay of LnmG with ACPs or PCP and [2-^{14}C]Malonyl CoA. LnmG-catalyzed loading of the malonyl group from malonyl CoA to ACPs or PCP was assayed in a two-step reaction. First, apo-ACPs or PCP was phosphopantetheinylated by *Streptomyces verticillus* phosphopatetheinyl transferase (Svp) with CoA (20, 21). A typical reaction of 75 μl contained 100 mM Tris·HCl, pH 7.5, 12.5 mM $MgCl_2$, 2.5 mM DTT, 33.3 μM CoA, 10 μM ACP or PCP, and 2 μM Svp incubated at 25°C for 60 min. Second, a mixture of 2 μM LnmG and 10 μl of [2-^{14}C]malonyl CoA [200 μM, 51 mCi/mmol (1 Ci = 37 GBq), Perkin–Elmer] in a 15-μl volume was added to each reaction. The reaction was incubated at 25°C and subsequently quenched by the addition of 900 μl of acetone at various time points. Proteins were precipitated by centrifugation at 4°C for 30 min after being frozen at −80°C for at least 1 h. The protein pellet was redissolved in 1× sampler buffer and separated on 4–15% SDS/PAGE gels (Bio-Rad). The resolved gels were visualized by Coomassie blue staining and phosphorimaging (LE phosphor screen, Amersham Pharmacia).

HPLC and Electrospray Ionization MS (ESI-MS) Analyses of Apo-, Holo-, and Malonyl-ACPs. For HPLC preparation of apo-, holo-, and malonyl-ACPs, the phosphopantetheinylation reaction or complete loading reaction was scaled up three times, and cold malonyl CoA was used instead. The second step of loading the malonyl group to holo-ACP proceeded for 10 min at 25°C. HPLC analysis was carried out on a Jupiter C-18 column (5 μm, 300 Å, 250 × 4.6 mm, Phenomenex, Belmont, CA), eluted with a gradient from 85% buffer A (H_2O + 0.1% CF_3CO_2H) to 90% buffer B (CH_3CN + 0.1% CF_3CO_2H) in 25 min at a flow rate of 1 ml/min, and UV detection was at 220 nm to separate the proteins in the reaction mixture. Individual protein peaks were collected, lyophilized, and redissolved in H_2O for ESI-MS analyses. The latter was performed on an Agilent (Palo Alto, CA) 1000 HPLC-MSD SL instrument.

LNM Yield Improvement by Overexpressing *lnmG* in *S. atroolivaceus* SB3004. To construct the *lnmG* overexpression constructs, a 450-bp *Eco*RI–*Sac*I fragment that harbors the *ErmE** fragment (18) and a 2,883-bp *Sac*I–*Bgl*II fragment of the *lnmG* gene were cloned into pBS3030 [a low-copy-number vector derived from the SCP2* origin of replicon (18)] and pBS3031 [a high-copy-number vector derived from the pIJ101 origin of replicon (18)] vectors, respectively, to yield pBS3032 and pBS3033. The latter was introduced into *S. atroolivaceus* SB3004 by conjugation, and the resultant SB3005 and SB3006 strains that harbor pBS3032 and pBS3033, respectively, were cultured and analyzed by HPLC for LNM production with the *S. atroolivaceus* S-140 wild-type strain as a control (17).

Results and Discussion

LnmI and LnmJ are "AT-less" PKSs That Are Essential for LNM Biosynthesis. We recently cloned and sequenced the *lnm* biosynthetic gene cluster from *S. atroolivaceus* S-140 and confirmed its involvement in LNM biosynthesis by genetic and biochemical characterization of the *lnmGHI* locus (17). We subsequently established the *lnm* gene cluster boundaries by sequential inactivation of ORFs from both ends of the sequenced 172-kb DNA region, defining the *lnm* gene cluster to be comprised of 27 ORFs (unpublished data). Among the genes identified from the *lnm* cluster are *lnmG*, *lnmI*, *lnmJ*, *lnmQ*, and *lnmP* that together encode the Lnm hybrid NRPS–PKS megasynthetase (Fig. 1).

LnmQ and LnmP are NRPS adenylation enzyme (22) and PCP (23), respectively, constituting the loading module. LnmI contains the previously characterized thiazole-forming NRPS module (17) as well as PKS module 3 and the KS domain of PKS module 4. LnmJ harbors PKS modules 4–8 plus a thioesterase domain. The Lnm megasynthetase-templated synthesis of **1** could be envisaged to begin at LnmQ and end with the cyclization of the full-length linear peptide–polyketide intermediate by the thioesterase domain of LnmJ to yield a macrolactam intermediate such as **2** (Fig. 1). Although it remains unclear what the origin of the 1,3-dioxo-1,2-dithiolane is and how it is spirofused to the 18-membered macrolactam ring, subsequent modification of **2** by the action of post-PKS enzymes could be envisaged to furnish **1** (Fig. 1).

The deduced Lnm NRPS and PKS functions are consistent with what would be required for the biosynthesis of **1** from the amino acid and acyl CoA precursors. However, the Lnm hybrid NRPS–PKS megasynthetase is characterized by several intrigu-

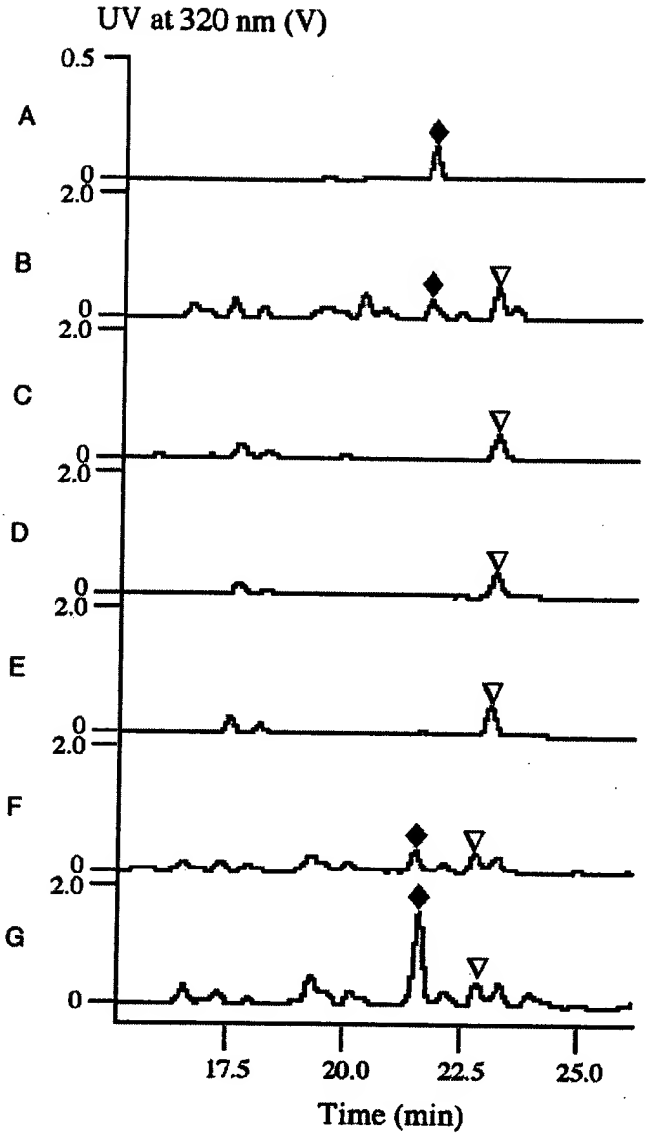

Fig. 2. HPLC analysis of LNM production by *S. atroolivaceus* wild-type and recombinant strains. (*A*) LNM standard. (*B*) S-140. (*C*) SB3002 (Δ*lnmI*). (*D*) SB3003 (Δ*lnmJ*). (*E*) SB3004 (Δ*lnmG*). (*F*) SB3005 (SB3004 harboring the *lnmG* overexpression plasmid of pBS3032). (*G*) SB3006 (SB3004 harboring the *lnmG* overexpression plasmid of pBS3033). ♦, LNM; ▽, an unknown metabolite with production that is independent to LNM biosynthesis.

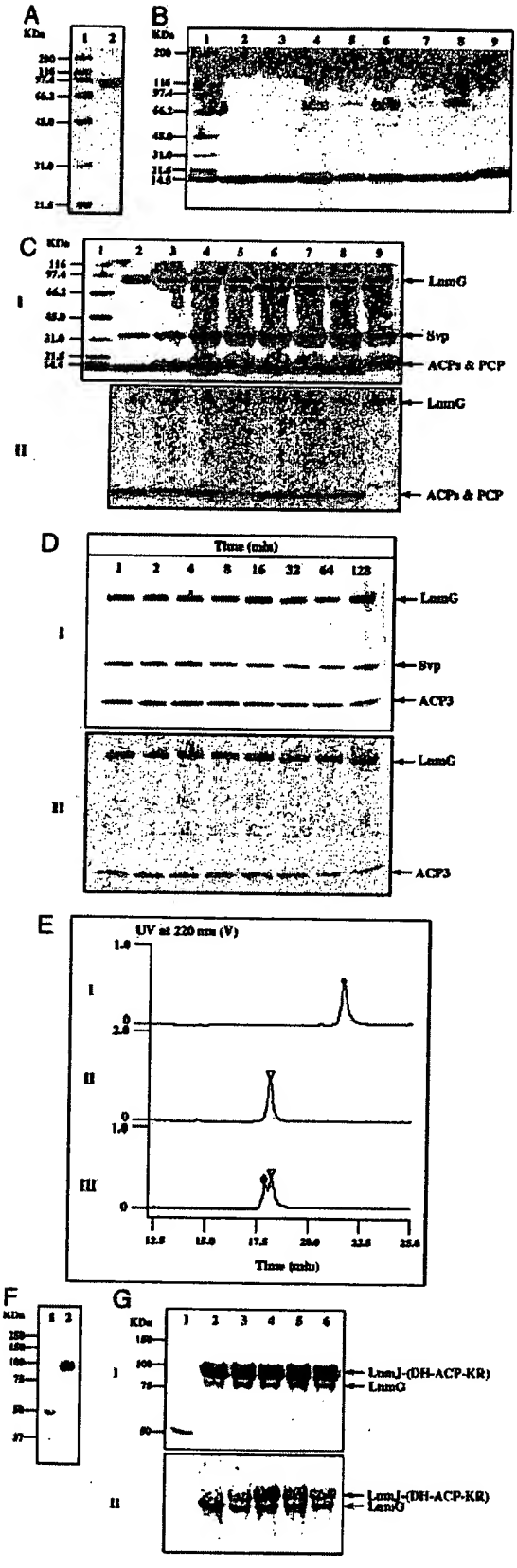

Fig. 3. *In vitro* assays of LnmG-catalyzed loading of malonyl CoA to individual LnmIJ PKS ACP domains and the LnmJ-(DH-ACP-KR) tridomain protein. (*A*) Purified LnmG on 4–15% SDS/PAGE. Lane 1, molecular mass standards; lane 2, LnmG. (*B*) Purified LnmIJ ACPs and LnmP on 4–15% SDS/PAGE. Lane 1, molecular mass standards; lane 2, ACP3; lane 3, ACP4; lane 4, ACP5; lane 5, ACP6-1; lane 6, ACP6-2; lane 7, ACP7; lane 8, ACP8; lane 9, LnmP. The numbers after the ACPs refer to the PKS modules from which they are derived with 6-1 and 6-2 to indicate the first and second ACPs, respectively, for PKS module 6. (*C*) Incubation of holo-ACPs or PCP with [2-^{14}C]malonyl CoA and LnmG as visualized on 4–15% SDS/PAGE (I) and by phosphorimaging (II). Lane 1, molecular mass standards; lanes 2–8; ACP3–ACP8; lane 9, LnmP. (*D*) Time course of LnmG-catalyzed loading of [2-^{14}C]malonyl CoA to ACP3 as visualized on 4–15% SDS/PAGE (I) and by phosphorimaging (II). (*E*) HPLC analysis of LnmG-catalyzed loading of malonyl CoA to ACP3. I, a negative control in the absence of Svp; II, a negative control in the absence of LnmG; III, complete assay. ●, apo-ACP3; ▽, holo-ACP3; ♦, malonyl-S-ACP3. (*F*) Purified LnmJ-(DH-ACP-KR) on 9% SDS/PAGE. Lane 1, molecular mass standards; lane 2, LnmJ-(DH-ACP-KR). (*G*) Incubation of holo-LnmJ-(DH-ACP-KR) with [2-^{14}C]malonyl CoA and LnmG as visualized on 9% SDS/PAGE (I) and by phosphorimaging (II). Lane 1, molecular mass standards; lane 2, a negative control in the absence of Svp; lanes 3–6, complete assay with incubation times of 2, 5, 15, and 60 min, respectively.

ing features, and most strikingly, it lacks the cognate AT domain from all six PKS modules, i.e., the LnmJ and LnmI are AT-less type I PKSs (Fig. 1). To investigate the role of *lnmI* and *lnmJ in vivo*, we replaced *lnmI* and *lnmJ* with mutant copies in which domains of NRPS module 2 and PKS module 3 (for *lnmI*) and of PKS modules 7 and 8 (for *lnmJ*) were substituted with the apramycin-resistance gene, *aac(3)IV*, respectively. The resultant *S. atroolivaceus* SB3002 and SB3003 mutant strains lost its ability to produce **1** (Fig. 2 *C* and *D*), confirming that *lnmI* and *lnmJ* are essential for **1** production.

LnmG Is a Discrete AT Protein That Is Essential for LNM Biosynthesis. Lack of AT domain in all six PKS modules raises the question of how the LnmI and LnmJ PKSs are charged with the extender unit malonyl CoA for the biosynthesis of **1**. To search for the missing AT activity, we reexamined genes within the *lnm* cluster and identified *lnmG*, the deduced product of which (the N-terminal half) showed high sequence homology to AT domains (3). We inactivated *lnmG* by replacing it with a mutant copy in which *lnmG* was disrupted by *aac(3)IV*. The resultant *S. atroolivaceus* SB3004 mutant lost its ability to produce **1**, confirming that *lnmG* is essential for **1** production (Fig. 2*E*). These results inspired us to propose that LnmG provides the AT activity in trans to LnmI and LnmJ and loads the malonyl CoA extender unit to all ACP domains of the six PKS modules for **1** biosynthesis (Fig. 1).

LnmG Loads Malonyl CoA in Trans to All Six PKS Modules of LnmI and LnmJ. To validate this hypothesis, we expressed both *lnmG* and the seven ACP domains from the six PKS modules encoded by *lnmI* and *lnmJ* as well as *lnmP* in *E. coli* and purified the resultant LnmG (Fig. 3*A*), ACPs, and LnmP PCP (Fig. 3*B*) as His_6-tagged fusion proteins to homogeneity. [We included LnmP, a discrete PCP identified within the *lnm* gene cluster (17), as a negative control to demonstrate that LnmG discriminates the LnmI and LnmJ PKS ACPs from other carrier proteins.] Because most ACPs or PCPs overproduced in *E. coli* are in the nonfunctional apo forms, we incubated them with CoA and the Svp phosphopantetheinyl transferase to ensure that all carrier proteins are converted into the functional holo forms (20, 21). We incubated the holo-carrier proteins with [2-^{14}C]malonyl CoA and LnmG to directly test malonyl CoA extender-unit loading. The reaction mixtures were subjected to (*i*) SDS/PAGE and phosphorimaging to detect specific loading of the [2-^{14}C]malonyl group to the phosphopantetheinyl group of ACPs (Fig. 3 *C* and *D*) and (*ii*) HPLC (Fig. 3*E*) and ESI-MS analyses to confirm the predicted molecular species (Table 1).

LnmG specifically and efficiently catalyzes the loading of the malonyl CoA extender unit to the LnmI and LnmJ PKS ACPs, and no loading was observed in the absence of LnmG. After 5

Table 1. ESI-MS analysis of apo-, holo-, and malonyl-S-ACPs

	apo-ACP [M + H]$^+$		holo-ACP [M + H]$^+$		Malonyl-S-ACP [M + H]$^+$	
ACPs*	Calcd.	Found	Calcd.	Found	Calcd.	Found
LnmI-ACP3	11,702	11,700	12,042	12,040	12,128	12,126
LnmJ-ACP4	12,245	12,241	12,585	12,582	12,671	12,669
LnmJ-ACP5	12,520	12,517	12,860	12,857	12,946	12,943
LnmJ-ACP6-1	12,209	12,206	12,549	12,546	12,635	12,632
LnmJ-ACP6-2	12,151	12,147	12,491	12,486	12,577	12,572
LnmJ-ACP-7	12,322	12,318	12,662	12,665	12,748	12,751
LnmJ-ACP-8	12,090	12,087	12,430	12,427	12,516	12,512

*The numbers after ACPs refer to the PKS modules from which they are derived with 6-1 and 6-2 to indicate the first and second ACPs, respectively, for PKS module 6.

min of incubation in the presence of LnmG and [2-^{14}C]malonyl CoA, six of the seven ACPs were loaded efficiently with the malonyl group (Fig. 3*C*, lanes 2–4 and 6–8) with ACP6-1 being loaded less efficiently (Fig. 3*C*, lane 5), and no loading was observed for the LnmP PCP (Fig. 3*C*, lane 9). [The observation that ACP6-1 is loaded less efficiently is consistent with the finding that PKS module 6 has two ACP domains, one of which may be the preferred site for malonyl CoA loading (Fig. 1).] LnmG was also labeled by [2-^{14}C]malonyl CoA (Fig. 3 *C*, *D*, and *G*). ATs are known to form acyl-*O*-enzyme intermediates at their active-site Ser residues before transferring the acyl groups from their CoA substrates to the nucleophilic recipients such as an ACP, and labeling of ATs by extender units has been confirmed for both the AT domain of type I PKS (24) and the malonyl CoA–ACP transacylase of type II PKS (25). Interestingly, the extent of ACP labeling decreased with longer incubation time, reaching a maximum in the first 5 min and falling to <10% after ≈2 h as exemplified by LnmI ACP3 (Fig. 3*D*). The malonyl-*S*-ACP product is apparently not stable under the assay condition, the malonyl group of which undergoes hydrolysis in the absence of chain elongation. Analogous results have been found for both the 6-deoxyerythronolide B synthase (24) and fatty acid synthase (26). In contrast, LnmG labeling appeared to be constant (Fig. 3*D*), suggesting that the malonyl-*O*-LnmG species is stable under the assay condition. To exclude any ambiguity associated with these assays, we subsequently purified the apo-, holo-, and malonyl-*S*-ACP species from the assay mixtures by HPLC, as exemplified by the LnmI–ACP3 (Fig. 3*E*), and established their identities by ESI-MS analysis. As summarized in Table 1, distinct [M + H]$^+$ ions at m/z values exact or near the calculated values of the corresponding ACP species were observed for all samples analyzed.

Finally, we expressed the PKS module 4 as a tridomain protein, LnmJ-(DH-ACP-KR), in *E. coli* and purified it as a His_6-tagged fusion protein (Fig. 3*F*). LnmJ-(DH-ACP-KR) was similarly assayed in the presence of LnmG and [2-^{14}C]malonyl CoA to confirm that LnmG can load the malonyl group to a multidomain PKS module as efficiently as to individual ACP domains. As summarized in Fig. 3*G*, LnmG efficiently loaded the malonyl group to holo-LnmJ-(DH-ACP-KR) (lanes 3–6); the extent of [2-^{14}C]malonyl labeling reached a maximum in 5 min (lane 4) and decreased with longer incubation time as a result of hydrolysis (lanes 4–6), and no loading was observed with the apo-LnmJ-(DH-ACP-KR) protein (lane 2).

The Interaction Among LnmG, LnmI, and LnmJ Represents a Previously Uncharacterized PKS Architecture. Among type I PKSs characterized to date, each module minimally contains the KS, AT, and ACP domains (2, 3). Limited structural studies suggested that deoxyerythronolide B synthase, the archetype of type I PKSs, forms a parallel homodimer, possibly a helical structure (27, 28). At the core of the helix is a tetrahedron formed by the KS and AT domains of each module with the ACP domain brought into close juxtaposition with the KS domain of the opposite subunit. The optional reductive domains form loops that protrude out from the central core while remaining within range of the phosphopantetheine arm of the adjacent ACP. This model, consistent with the recently solved crystal structure of the homodimer of the thioesterase domain of deoxyerythronolide B synthase (29), has important general implications for both the mechanism of enzyme catalysis and genetic engineering of type I PKSs.

We now report that the LnmI and LnmJ are AT-less type I PKSs consisting minimally of the KS and ACP domains and that LnmG is a discrete AT enzyme that interacts with LnmI and LnmJ to form functional type I PKS modules by providing the AT activity in trans. These findings suggest an alternative model

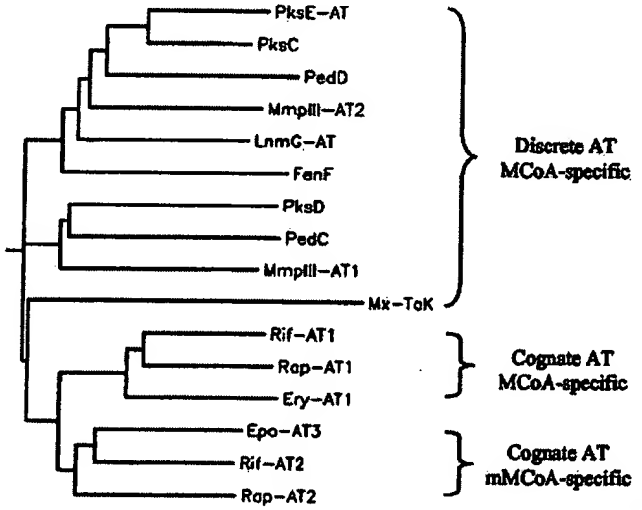

Fig. 4. Phylogenetic analysis of LnmG AT and its homologs from other AT-less type I PKS clusters and their relationships to cognate ATs from type I PKS clusters. The LnmG AT and its homologs fall into distinct groups that differ from cognate ATs of type I PKSs. Multiple sequence alignment and phylogenetic analysis were performed by the GCG program. We predict that PksC/PksD/PksE-AT [for the unknown polyketide in *Bacillus subtilis* (33)], PedC/PedD [for pederin in a bacterial symbiont of *Paederus* beetles (30)], MmpIII-AT1/AT2 [for mupirocin in *Pseudomonas fluorescens* (GenBank accession no. AF318063)], FenF [for mycosubtilin in *B. subtilis* ATCC6633 (31)], and Mx-TaK [for Ta1 in *Myxococcus xanthus* (32)], acting in a mechanistic analogy to LnmG, load the malonyl CoA extender unit onto the AT-less PKS modules in trans for polyketide biosynthesis in these clusters. For cognate ATs from rifamycin (Rif), rapamycin (Rap), erythromycin (Ery), and epothilone (Epo) clusters, protein GenBank accession numbers are given after the protein names: Rif-AT4 and Rif-AT2, AF04570; Rap-AT1 and Rap-AT2, X86780; Ery-AT1, Q03131; Epo-AT3, AF217189. MCoA, malonyl CoA; mMCoA, methyl malonyl CoA.

for type I PKS in which the KS and ACP domains of each module could minimally constitute the core structure. Although AT-less PKS is very rare and the recently reported putative pederin cluster was the only other example to our knowledge that lacks the cognate AT domain from all modules of a type I PKS (30), individual modules that lack the AT domain have been noted in a few other type I PKS or hybrid NRPS–PKS systems (31–35). The AT activity in all these AT-less type I PKS clusters remains unknown. We propose that these AT-less PKS clusters represent a previously unknown architecture for type I PKSs, which is characterized by discrete, iteratively acting AT(s), that loads the extender units in trans to AT-less type I PKS(s). Inspired by this model, we performed sequence analysis of all AT-less type I PKS clusters known to date to search for the discrete AT enzymes. Preliminary phylogenetic analysis indeed has led to the identification of LnmG homologs from these clusters (Fig. 4), which, in a mechanistic analogy to LnmG, could provide the AT activity in trans to the AT-less PKS modules for polyketide biosynthesis.

The AT-Less Type I PKS Architecture Provides Opportunities for PKS Engineering. The model of AT-less type I PKS reported here provides opportunities for PKS engineering. For example, given the mechanism that LnmG is responsible for the loading of the malonyl CoA extender unit to all six PKS modules of the Lnm hybrid NRPS–PKS megasynthetase, we reasoned that LnmG could be a rate-limiting factor for **1** biosynthesis. We therefore explored yield improvement for **1** by overexpressing *lnmG* under the constitutive *ErmE** promoter in both low- and high-copy-number vectors in *S. atroolivaceus* SB3004. The resultant recombinant strains SB3005, harboring the low-copy-number expression construct pBS3032, and SB3006, harboring the high-copy-number expression construct pBS3033, produce 3- to 5-fold more **1** (Fig. 2 *F* and *G*) than the wild-type S-140 strain (Fig. 2*B*) as determined by HPLC analysis. We also envisage applying *lnmG* and its homologs to either type I PKSs or AT-less PKSs to alter the extender-unit specificity by combinatorial biosynthesis methods, further expanding the size and diversity of polyketide natural products.

We thank Kyowa Hakko Kogyo (Tokyo) for an authentic sample of LNM, the wild-type *S. atroolivaceus* S-140 strain, and assistance in sequencing the *lnm* gene cluster. B.S. is a recipient of National Science Foundation Career Award MCB9733938 and National Institutes of Health Independent Scientist Award AI51689. This work was supported in part by University of California BioSTAR Program Grant Bi099-10045 and KOSAN Biosciences (Hayward, CA).

1. O'Hagan, D. (1991) *Polyketides* (Horwood, Chichester, U.K.).
2. Hopwood, D. A. (1997) *Chem. Rev. (Washington, DC)* **97**, 2465–2497.
3. Staunton, J. & Weissman, K. J. (2001) *Nat. Prod. Rep.* **18**, 380–416.
4. Shen, B. (2000) *Top. Curr. Chem.* **209**, 1–51.
5. McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. (1993) *Science* **262**, 1546–1550.
6. Bisang, C., Long, P. F., Cortes, J., Westcott, J., Crosby, J., Matharu, A., Cox, R. J., Simpson, T. J., Staunton, J. & Leadlay, P. F. (1999) *Nature* **401**, 502–505.
7. Funa, N., Ohnishi, Y., Fujii, I., Shibuya, M., Ebizuka, Y. & Horinouchi, S. (1999) *Nature* **400**, 897–899.
8. Moore, B. S. & Hopke, J. N. (2001) *Chembiochem* **2**, 35–38.
9. Cane, D. E., Walsh, C. T. & Khosla, C. (1998) *Science* **282**, 63–68.
10. Cane, D. E. & Walsh, C. T. (1999) *Chem. Biol.* **6**, R319–R325.
11. Du, L. & Shen, B. (2001) *Curr. Opin. Drug. Discov. Dev.* **4**, 215–228.
12. Walsh, C. T. (2002) *Chembiochem* **3**, 124–134.
13. Hara, M., Asano, K., Kawamoto, I., Takiguchi, T., Katsumata, S., Takahashi, K. & Nakano, H. (1989) *J. Antibiot.* **42**, 1768–1774.
14. Hirayama, N. & Matsuzawa, E. S. (1993) *Chem. Lett.*, 1957–1958.
15. Kanda, Y. & Fukuyama, T. (1993) *J. Am. Chem. Soc.* **115**, 8451–8452.
16. Nakano, H. & Tamaoki, T. (1992) in *Harnessing Biotechnology for the 21st Century: Proceedings of the Ninth International Biotechnology Symposium and Exposition, Crystal City, Virginia*, eds. Ladisch, M. R. & Bose, A. (Am. Chem. Soc., Washington, DC), pp. 72–75.
17. Cheng, Y.-Q., Tang, G.-L. & Shen, B. (2002) *J. Bacteriol.* **184**, 7013–7024.
18. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000) *Practical Streptomyces Genetics* (The John Innes Foundation, Norwich, U.K.).
19. Bierman, M., Logan, R., O'Brien, K., Seno, E. T., Nagaraja, R. & Schoner, B. E. (1992) *Gene* **116**, 43–49.
20. Walsh, C. T., Gehring, A. M., Weinreb, P. H., Quadri, L. E. N. & Flugel, R. S. (1997) *Curr. Opin. Chem. Biol.* **1**, 309–315.
21. Sanchez, C., Du, L., Edwards, D. J., Toney, M. D. & Shen, B. (2001) *Chem. Biol.* **8**, 725–738.
22. Conti, E., Stachelhaus, T., Marahiel, M. A. & Brick, P. (1997) *EMBO J.* **16**, 4147–4183.
23. Weber, T., Baumgartner, R., Renner, C., Marahiel, M. A. & Holak, T. A. (2000) *Structure Fold. Des.* **8**, 407–418.
24. Marsden, A. F. A., Caffrey, P., Aparicio, J. F., Loughran, M. S., Staunton, J. & Leadlay, P. F. (1994) *Science* **263**, 378–380.
25. Carreras, C. W. & Khosla, C. (1998) *Biochemistry* **37**, 2084–2088.
26. Wakil, S. J. (1989) *Biochemistry* **28**, 4523–4530.
27. Staunton, J., Caffrey, P., Aparicio, J. F., Roberts, G., Bethell, S. & Leadlay, P. F. (1996) *Nat. Struct. Biol.* **3**, 188–192.
28. Gokhale, R. S., Lau, J., Cane, D. E. & Khosla, C. (1998) *Biochemistry* **37**, 2524–2528.
29. Tsai, S.-C., Miercke, L. J. W., Krucinski, J., Gokhale, R., Chen, J. C.-H., Foster, P. G., Cane, D. E., Khosla, C. & Stroud, R. M. (2001) *Proc. Natl. Acad. Sci. USA* **99**, 14002–14007.
30. Piel, J. (2002) *Proc. Natl. Acad. Sci. USA* **98**, 14808–14813.
31. Duitman, E. H., Hamoen, L. W., Rembold, M., Venema, G., Seitz, H., Saenger, W., Bernhard, F., Reinhardt, R., Schmidt, M., Ullrich, C., *et al.* (1999) *Proc. Natl. Acad. Sci. USA* **96**, 13294–13299.
32. Paitan, Y., Alon, G., Orr, E., Ron, E. Z. & Rosenberg, E. (1999) *J. Mol. Biol.* **286**, 465–474.
33. Albertini, A. M., Caramori, T., Scoffone, F., Scotti, C. & Galizzi, A. (1995) *Microbiology* **141**, 299–309.
34. Huang, G., Zhang, L. & Birch, R. G. (2001) *Microbiology* **147**, 631–642.
35. Zhu, G., LaGier, M. J., Stejskal, F., Millership, J. J., Cai, X. & Keithly, J. S. (2002) *Gene* **298**, 79–89.

1: Q03131. Erythronolide syn...[gi:416965] BLink, Domains, Links

```
LOCUS       Q03131                  3491 aa            linear   BCT 15-JUN-2002
DEFINITION  Erythronolide synthase, modules 1 and 2 (ORF 1)
            (6-deoxyerythronolide B synthase I) (DEBS 1).
ACCESSION   Q03131
VERSION     Q03131  GI:416965
DBSOURCE    swissprot: locus ERY1_SACER, accession Q03131;
            class: standard.
            created: Oct 1, 1993.
            sequence updated: Oct 1, 1993.
            annotation updated: Jun 15, 2002.
            xrefs: gi: 152691, gi: 7264827, gi: 294869, gi: 294871
            xrefs (non-sequence databases): HSSPP25715, InterProIPR002198,
            InterProIPR001227, InterProIPR000794, InterProIPR003880,
            PfamPF00106, PfamPF00109, PfamPF00550, PfamPF00698, PfamPF02801,
            PROSITEPS00012, PROSITEPS00606, PROSITEPS50075
KEYWORDS    Transferase; Acyltransferase; Antibiotic biosynthesis; NADP;
            Repeat; Phosphopantetheine; Multifunctional enzyme.
SOURCE      Saccharopolyspora erythraea
  ORGANISM  Saccharopolyspora erythraea
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Pseudonocardineae; Pseudonocardiaceae; Saccharopolyspora.
REFERENCE   1  (residues 1 to 3491)
  AUTHORS   Donadio,S., Staver,M.J., McAlpine,J.B., Swanson,S.J. and Katz,L.
  TITLE     Modular organization of genes required for complex polyketide
            biosynthesis
  JOURNAL   Science 252 (5006), 675-679 (1991)
  MEDLINE   91220065
   PUBMED   2024119
  REMARK    SEQUENCE FROM N.A.
REFERENCE   2  (residues 1 to 3491)
  AUTHORS   Donadio,S. and Staver,M.J.
  TITLE     IS1136, an insertion element in the erythromycin gene cluster of
            Saccharopolyspora erythraea
  JOURNAL   Gene 126 (1), 147-151 (1993)
  MEDLINE   93231529
   PUBMED   8386127
  REMARK    SEQUENCE OF 3474-3491 FROM N.A.
COMMENT     ---------------------------------------------------------------
            This SWISS-PROT entry is copyright. It is produced through a
            collaboration between the Swiss Institute of Bioinformatics and
            the EMBL outstation - the European Bioinformatics Institute.
            The original entry is available from http://www.expasy.ch/sprot
            and http://www.ebi.ac.uk/sprot
            ---------------------------------------------------------------.
            [CATALYTIC ACTIVITY] 6 malonyl-CoA + propionyl-CoA = 7 CoA +
            6-deoxyerythronolide B.
            [COFACTOR] NADP; CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTETHEINES.
```

[PATHWAY] COMPLEX POLYKETIDE FORMATION IN ERYTHROMYCIN BIOSYNTHESIS.
[MISCELLANEOUS] IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX FAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.
[MISCELLANEOUS] BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR), DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.
[SIMILARITY] TO FATTY ACID SYNTHASE (FAS).
[SIMILARITY] CONTAINS 3 ACYL CARRIER DOMAINS.

FEATURES	Location/Qualifiers
source	1..3491 /organism="Saccharopolyspora erythraea" /db_xref="taxon:1836"
gene	1..3491 /gene="ERYA"
Protein	1..3491 /gene="ERYA" /product="Erythronolide synthase, modules 1 and 2" /EC_number="2.3.1.94"
Region	1..1972 /gene="ERYA" /region_name="Domain" /note="MODULE 1."
Region	1..375 /gene="ERYA" /region_name="Domain" /note="ACYLTRANSFERASE (AT) 1."
Site	145 /gene="ERYA" /site_type="active" /note="ACYL-ESTER INTERMEDIATE."
Region	414..484 /gene="ERYA" /region_name="Domain" /note="ACYL CARRIER (ACP) 1."
Site	447 /gene="ERYA" /site_type="binding" /note="PHOSPHOPANTETHEINE (BY SIMILARITY)."
Region	503..961 /gene="ERYA" /region_name="Domain" /note="BETA-KETOACYL SYNTHASE 1."
Site	677 /gene="ERYA" /site_type="active" /note="BETA-KETOACYL SYNTHASE."
Region	1030..1356 /gene="ERYA" /region_name="Domain" /note="ACYLTRANSFERASE (AT) 2."
Site	1128

```
                     /gene="ERYA"
                     /site_type="active"
                     /note="ACYL-ESTER INTERMEDIATE."
     Region          1611..1794
                     /gene="ERYA"
                     /region_name="Domain"
                     /note="BETA-KETOACYL REDUCTASE 1."
     Site            1614..1660
                     /gene="ERYA"
                     /site_type="np-binding"
                     /note="NADP."
     Region          1888..1958
                     /gene="ERYA"
                     /region_name="Domain"
                     /note="ACYL CARRIER (ACP) 2."
     Site            1921
                     /gene="ERYA"
                     /site_type="binding"
                     /note="PHOSPHOPANTETHEINE (BY SIMILARITY)."
     Region          1979..3491
                     /gene="ERYA"
                     /region_name="Domain"
                     /note="MODULE 2."
     Region          1979..2441
                     /gene="ERYA"
                     /region_name="Domain"
                     /note="BETA-KETOACYL SYNTHASE 2."
     Site            2148
                     /gene="ERYA"
                     /site_type="active"
                     /note="BETA-KETOACYL SYNTHASE."
     Region          2507..2854
                     /gene="ERYA"
                     /region_name="Domain"
                     /note="ACYLTRANSFERASE (AT) 3."
     Site            2598
                     /gene="ERYA"
                     /site_type="active"
                     /note="ACYL-ESTER INTERMEDIATE."
     Region          3055..3237
                     /gene="ERYA"
                     /region_name="Domain"
                     /note="BETA-KETOACYL REDUCTASE 2."
     Site            3058..3104
                     /gene="ERYA"
                     /site_type="np-binding"
                     /note="NADP."
     Region          3334..3404
                     /gene="ERYA"
                     /region_name="Domain"
                     /note="ACYL CARRIER (ACP) 3."
     Site            3367
                     /gene="ERYA"
                     /site_type="binding"
                     /note="PHOSPHOPANTETHEINE (BY SIMILARITY)."
ORIGIN      AT
        1 msgprsrtts rrtpvrigav vvasstsell dglaavadgr phasvvrgva rpsapvvfvf
       61 pgqgaqwagm agellgesrv faaamdacar afepvtdwtl aqvldspeqs rrvevvqpal
      121 favqtslaal wrsfgvtpda vvghsigela aahvcgaaga adaaraaalw sremiplvgn
      181 gdmaavalsa deiepriarw dddvvlagvn gprsvlltgs pepvarrvqe lsaegvraqv
```

```
 241 invsmaahsa qvddiaegmr salawfapgg sevpfyaslt ggavdtrelv adywrrsfrl
 301 pvrfdeairs alevgpgtfv easphpvlaa alqqtldaeg ssaavvptlq rgqqgmrrfl
 361 laaaqaftgg vavdwtaayd dvgpnpalcr ssrrprrkts rpspastgtr hrtccerlla
 421 vvngetaala greadaeatf relgldsvla aqlrakvsaa igrevniall ydhptprala
 481 ealaagteva qretrartne aapgepvavv amacrlpggv stpeefwell segrdavagl
 541 ptdrgwdlds lfhpdptrsg tahqrgggfl teatafdpaf fgmspreala vdpqqrlmle
 601 lswevlerag ipptslqasp tgvfvglipq eygprlaegg egvegylmtg tttsvasgri
 661 aytlglegpa isvdtacsss lvavhlacqs lrrgesslam aggvtvmptp gmlvdfsrmn
 721 slapdgrcka fsagangfgm aegagmllle rlsdarrngh pvlavlrgta vnsdgasngl
 781 sapngraqvr viqqalaesg lgpadidave ahgtgtrlgd piearalfea ygrdreqplh
 841 lgsvksnlgh tqaaagvagv ikmvlamrag tlprtlhase rskeidwssg aislldepep
 901 wpagarprra gvssfgisgt nahaiieeap qvvegervea gdvvapwvls assaeglraq
 961 aarlaahlre hpgqdprdia yslatgraal phraafapvd esaalrvldg latgnadgaa
1021 vgtsraqqra vfvfpgqgwq wagmavdlld tspvfaaalr ecadalephl dfevipflra
1081 eaarreqdaa lstervdvvq pvmfavmvsl asmwrahgve paavighsqg eiaaacvaga
1141 lslddaarvv alrsrviatm pgnkgmasia apagevrari gdrveiaavn gprsvvvagd
1201 sdeldrlvas cttecirakr lavdyashss hvetirdalh aelgedfhpl pgfvpffstv
1261 tgrwtqpdel dagywyrnlr rtvrfadavr alaeqgyrtf levsahpilt aaieeigdgs
1321 gadlsaihsl rrgdgsladf gealsrafaa gvavdwesvh lgtgarrvpl ptypfqrerv
1381 wlepkpvarr stevdevsal ryriewrptg ageparldgt wlvakyagta detstaarea
1441 lesagarvre lvvdarcgrd elaerlrsvg evagvlslla vdeaepeeap lalasladtl
1501 slvqamvsae lgcplwtvte savatgpfer vrnaahgalw gvgrvialen pavwgglvdv
1561 pagsvaelar hlaavvsgga gedqlalrad gvygrrwvra aapatddewk ptgtvlvtgg
1621 tggvggqiar wlarrgaphl llvsrsgpda dgagelvael ealgarttva acdvtdresv
1681 rellggigdd vplsavfhaa atlddgtvdt ltgerieras rakvlgarnl heltreldlt
1741 afvlfssfas afgapglggy apgnayldgl aqqrrsdglp atavawgtwa gsgmaegava
1801 drfrrhgvie mppetacral qnaldraeyc pividvrwdr fllaytaqrp trlfdeidda
1861 rraapqapae prvgalaslp apereealfe lvrshaaavl ghasaervpa dqafaelgvd
1921 slsalelrnr lgaatgvrlp tttvfdhpdv rtlaahlaae lggatgaeqa apattapvde
1981 piaivgmacr lpgevdsper lwelitsgrd saaevpddrg wvpdelmasd aagtrahgnf
2041 magagdfdaa ffgispreal amdpqqrqal ettwealesa gippetlrgs dtgvfvgmsh
2101 qgyatgrprp edgvdgyllt gntasvasgr iayvlglegp altvdtacss slvalhtacg
2161 slrdgdcgla vaggvsvmag pevftefsrq galspdgrck pfsdeadgfg lgegsafvvl
2221 qrlsdarreg rrvlgvvags avnqdgasng lsapsgvaqq rvirrawara gitgadvavv
2281 eahgtgtrlg dpveasalla tygksrgssg pvllgsvksn ighaqaaagv agvikvllgl
2341 ergvvppmlc rgersglidw ssgeieladg vrewspaadg vrragvsafg vsgtnahvii
2401 aeppepepvp qprrmlpatg vvpvvlsart gaalraqagr ladhlaahpg iapadvswtm
2461 ararqhfeer aavlaadtae avhrlravad gavvpgvvtg sasdggsvfv fpgqgaqweg
2521 marellpvpv faesiaecda vlsevagfsv sevleprpda pslervdvvq pvlfavmvsl
2581 arlwracgav psavighsqg eiaaavvaga lsledgmrvv arrsravrav agrgsmlsvr
2641 ggrsdvekll addswtgrle vaavngpdav vvagdaqaar efleycegvg iraraipvdy
2701 ashtahvepv rdelvqalag itprraevpf fstltgdfld gteldagywy rnlrhpvefh
2761 savqaltdqg yatfievsph pvlassvqet lddaesdaav lgtlerdagd adrfltalad
2821 ahtrgvavdw eavlgraglv dlpgypfqgk rfwllpdrtt prdeldgwfy rvdwtevprs
2881 epaalrgrwl vvvpegheed gwtvevrsal aeagaepevt rgvgglvgdc agvvsllale
2941 gdgavqtlvl vreldaegid aplwtvtfga vdagspvarp dqaklwglgq vaslergprw
3001 tglvdlphmp dpelrgrlta vlagsedqva vradavrarr lspahvtats eyavpggtil
3061 vtggtaglga evarwlagrg aehlalvsrr gpdtegvgdl taeltrlgar vsvhacdvss
3121 repvrelvhg lieqgdvvrg vvhaaglpqq vaindmdeaa fdevvaakag gavhldelcs
3181 daelfllfss gagvwgsarq gayaagnafl dafarhrrgr glpatsvawg lwaaggmtgd
3241 eeavsflrer gvrampvpra laaldrvlas getavvvtdv dwpafaesyt aarprplldr
3301 ivttapsera gepeteslrd rlaglpraer taelvrlvrt statvlghdd pkavrattpf
3361 kelgfdslaa vrlrnllnaa tglrlpstlv fdhpnasava gfldaelgte vrgeapsala
3421 gldalegalp evpatereel vqrlermlaa lrpvaqaada sgtganpsgd dlgeagvdel
3481 lealgreldg d
```

//

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Oct. 2001, p. 4531–4537
0099-2240/01/$04.00+0 DOI: 10.1128/AEM.67.10.4531–4537.2001

Vol. 67, No. 10

Evidence for the Biosynthesis of Bryostatins by the Bacterial Symbiont "*Candidatus* Endobugula sertula" of the Bryozoan *Bugula neritina*

S. K. DAVIDSON,[1,2]† S. W. ALLEN,[3]‡ G. E. LIM,[1,2] C. M. ANDERSON,[1,2] AND M. G. HAYGOOD[1,2]*

Scripps Institution of Oceanography, Marine Biology Research Division, Center for Marine Biomedicine and Biotechnology,[1] and UCSD Cancer Center,[2] University of California, San Diego, La Jolla, California 92093-0202, and CalBioMarine Technologies, Inc., Carlsbad, California 92009[3]

Received 23 March 2001/Accepted 16 July 2001

The marine bryozoan, *Bugula neritina*, is the source of the bryostatins, a family of macrocyclic lactones with anticancer activity. Bryostatins have long been suspected to be bacterial products. *B. neritina* harbors the uncultivated gamma proteobacterial symbiont "*Candidatus* Endobugula sertula." In this work several lines of evidence are presented that show that the symbiont is the most likely source of bryostatins. Bryostatins are complex polyketides similar to bacterial secondary metabolites synthesized by modular type I polyketide synthases (PKS-I). PKS-I gene fragments were cloned from DNA extracted from the *B. neritina*-"*E. sertula*" association, and then primers specific to one of these clones, KSa, were shown to amplify the KSa gene specifically and universally from total *B. neritina* DNA. In addition, a KSa RNA probe was shown to bind specifically to the symbiotic bacteria located in the pallial sinus of the larvae of *B. neritina* and not to *B. neritina* cells or to other bacteria. Finally, *B. neritina* colonies grown in the laboratory were treated with antibiotics to reduce the numbers of bacterial symbionts. Decreased symbiont levels resulted in the reduction of the KSa signal as well as the bryostatin content. These data provide evidence that the symbiont *E. sertula* has the genetic potential to make bryostatins and is necessary in full complement for the host bryozoan to produce normal levels of bryostatins. This study demonstrates that it may be possible to clone bryostatin genes from *B. neritina* directly and use these to produce bryostatins in heterologous host bacteria.

Since 1968, bryostatins in extracts from the marine bryozoan *Bugula neritina* have been studied for their cellular activity and therapeutic potential (24). *B. neritina* is the sole source of this unique family of macrolides. The bryostatins are complex polyketides based on the bryopyran ring system that are similar to bacterial secondary metabolites (25) (Fig. 1). The bryostatins, which do not originate in the diet (J. Thompson and D. Mendola, Aquaculture production of bryostatin 1: nonconfidential results of the phase 1 project, CalBioMarine Technologies, Inc.), have been proposed to be products of bacteria associated with *B. neritina* (2). This bryozoan harbors an uncultivated gamma proteobacterium, "*Candidatus* Endobugula sertula," which is passed to the larvae prior to their release from the adult (15, 34). Bryostatins show excellent potential as therapeutic agents that act through protein kinase C (PKC) signal transduction to alter cellular activity (18, 30, 31). Phase II clinical trials of bryostatin 1 are under way for the treatment of leukemias, lymphomas, melanomas, and solid tumors (26). Bryostatin 1 also shows promise for treatment of ovarian and breast cancers and for enhancing lymphocyte survival during radiation treatment (3, 4, 5, 12, 13, 16, 20, 21, 28, 32, 33). Other bryostatins may prove to be valuable therapeutic agents as well (19). Research and development of the bryostatins are currently severely limited by inadequate availability of these compounds. If a bacterial source of bryostatins could be identified, new avenues for solving the supply problem may open, either through efforts to clone the bryostatin biosynthetic genes and then express them in a heterologous host or by culture of the bacterial source itself.

In this article we describe evidence that the bacterial symbiont "*E. sertula*" is involved in the biosynthesis of bryostatins. Previously, we have shown that a difference in the types of bryostatins found in *B. neritina* correlated with a genetic difference (by 16S ribosomal gene sequences) in the bacterial symbiont (7). However, the chemistry differences noted in that study could be explained by a difference in the host or another unidentified bacterial symbiont. Determining the source of bryostatins in the bryozoan bacterial association presents a difficult challenge due to the inability to culture the bacterial symbiont. Earlier attempts to separate the bacterial cells from the host as a way to determine the bryostatin source proved inconclusive (6). To investigate the link between the bacterial symbiont "*E. sertula*" and the production of bryostatins we applied molecular techniques that do not require the culture of the bacteria. In bacteria, complex polyketides are synthesized by modular type I polyketide synthases (PKS-I). These enzymes have a set of catalytic domains, including a β-ketoacyl synthase (KS) domain, for each elongation and modification step in the synthesis of the chain. In this study, the bacterial PKS-I gene fragments from KS domains were targeted for cloning and sequencing. The objectives of the research presented here were to show (i) that the genes coding for a type I complex polyketide synthase pathway are present in the DNA obtained from the symbiotic association, (ii) that genes required in the pathway, the KS genes, are found in the symbiont "*E. sertula*," and (iii) that a reduction of this symbiont results in a reduction of this detected gene and a reduction in bryostatin synthesis.

* Corresponding author. Mailing address: Scripps Institution of Oceanography, Marine Biology Research Division, 0202, University of California, San Diego, La Jolla, CA 92093-0202. Phone: (858) 534-5987. Fax: (858) 534-7313. E-mail: mhaygood@ucsd.edu.

† Present address: Civil and Environmental Engineering, University of Washington, Seattle, WA 98195-2700.

‡ Present address: Puracyp LLC, San Diego, CA 92126.

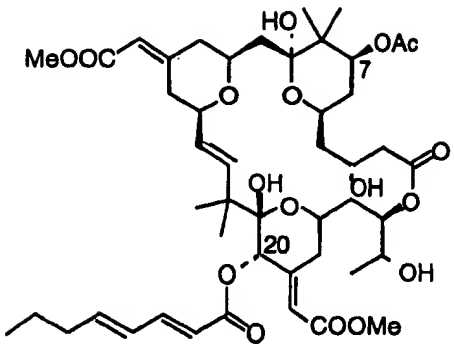

FIG. 1. Bryostatin 1 from *B. neritina*.

MATERIALS AND METHODS

DNA extraction. DNA for PCR amplification of PKS-I genes using degenerate KS primers was extracted from adult *B. neritina* using a modification of the method of Shure et al. (29). DNA extracted from *B. neritina* also contains bacterial DNA, which includes that of "*E. sertula*" and other environmental bacteria fouling the surface of the bryozoan. Adult *B. neritina* (2 g) was pulverized on dry ice, and then 6 ml of extraction buffer (8 M urea, 0.35 M NaCl, 0.05 M Tris-HCl [pH 7.5], 0.02 M EDTA, 2% sarcosyl) was added, followed by a phenol-choroform extraction, DNA precipitation, and 70% ethanol wash. The DNA was dissolved in a solution of 10 mM Tris-HCl (pH 7.5) and 10 mM EDTA and then passed through a Sephadex G-200 spin column (22) to remove PCR inhibitors.

DNA samples for amplification of PKS-I genes from larvae and for gene surveys of *B. neritina* populations with specific KSa primers were extracted from clean larvae and small portions of adult colonies as previously described, using a QIAamp tissue kit (Qiagen, Inc., Valencia, Calif.) (7, 15). DNA extracted from larvae contains primarily *B. neritina* and "*E. sertula*" DNA, with only slight contamination of DNA from other environmental bacteria.

Amplification, cloning, and sequencing of the KS domain DNA from *B. neritina*. Amplification of KS genes from *B. neritina* DNA was accomplished using the degenerate primers KSD1F and KSD1R (Table 1) designed from conserved regions of KS domains of bacterial PKS-I genes. Bacterial and fungal polyketide synthase and rat fatty acid synthase amino acid sequences (accession numbers M76767, P40872, Q03131, Q03132, Q03133, S41729, S43048, U24241, U31329, and U00023) were aligned and examined for conserved regions. The region targeted for amplification is downstream from the active-site cysteine. The forward primer is based on the motif HGTGT, which is conserved, but not specific, as it appears in fatty acid synthase (FAS) as well. The downstream primer is based on the motif GTNAHV, which is conserved and specific to bacterial PKS. The primers were degenerate with inosine used for positions with fourfold redundancy. PCRs contained approximately 1 ng of *B. neritina* DNA (either adult or larval) μl^{-1}, a 1 μM concentration of each primer, *Taq* polymerase, and buffer (Boehringer Mannheim Corp., Indianapolis, Ind.). To optimize annealing of degenerate KS primers, the annealing temperature was decreased with each temperature cycle from 60 to 40°C at a rate of 2°C per cycle (11 cycles) and then maintained at 40°C (39 cycles). Cycles included denaturation (94°C; 1 min), annealing, and extension (72°C; 1 min). PCR products of approximately 300 bp were cloned using a TOPO TA Cloning kit into Invitrogen pCR 2.1-TOPO vector as described by the manufacturer (Invitrogen Corp., Carlsbad, Calif.). Recombinant clones containing inserts were sequenced and analyzed. Primers specific to an abundant clone sequence from larval DNA, KSa, were designed for specific amplification of this gene fragment (Table 1 [Ksa1F and Ksa1R]). For the KSa-specific primers, the annealing temperature was 60°C, with other conditions the same (30 cycles).

Plasmid DNA of KS clones was prepared for sequencing using the QIAprep Spin Miniprep Kit (Qiagen, Inc.). All sequencing reactions were performed using the Applied Biosystems, Inc., PRISM Ready Reaction Dye Deoxy (or BigDye) terminator cycle sequencing kit as recommended by the manufacturer. Reactions were analyzed on an ABI automated sequencer (model 373A). Cloned genes were sequenced using primers directed against the cloning vector, pCR 2.1-TOPO (Invitrogen). Sequences were assembled using Sequencher and analyzed with BLAST (1) and by alignment with PKS and FAS sequences.

Design of mRNA KSa probe and in situ hybridizations. To determine if KSa is expressed in "*E. sertula*," in situ hybridizations of a KSa-specific RNA antisense probe were performed on larval *B. neritina*. Briefly, 300-bp fragments from a KSa clone were amplified by PCR using the primer pairs Ksa1F-T7Ksa1R and Ksa1R-T7Ksa1F to obtain antisense and sense strands respectively (8). T7Ksa1R and T7Ksa1F are primers with a T7 promoter at the 5′ ends. PCR products were purified using a QiaQuick PCR cleanup kit (Qiagen. Inc.) and used as templates in a T7 transcription reaction using the AmpliScribe T7 high-yield transcription kit as described by the manufacturer (EpiCentre Technologies, Inc., Madison, Wis.) with the addition of biotin-16-UTP (Boehringer Mannheim, Corp.) to obtain biotinylated probes. Another transcription was carried out using fluorescein-12-UTP to generate fluorescein-labeled probes for in situ hybridizations in *Escherichia coli*. Fluorescein probes were used because endogenous peroxidases in *E. coli* interfered with the avidin-horseradish peroxidase detection system used in the larvae.

In situ hybridizations in whole larvae and *E. coli* bacterial cells were performed as described by Haygood and Davidson (15). The annealing temperature for the 300-bp RNA probe was 47°C, and wash temperatures were 55, 37°C, and room temperature. In situ hybridizations of the KSa mRNA probe, sense and antisense, were performed with *E. coli* DH5 to control for probe binding to bacterial FAS.

Settlement and antibiotic treatment of *B. neritina* in the laboratory. Adult *B. neritina* was collected from Torrey Pines artificial reef (La Jolla, Calif., September 1996 and June 1998) and Palos Verdes (Long Beach, Calif., March 1997 and June 1998) and maintained in aquaria. Larvae were released into the aquaria, collected, rinsed in filtered (0.2-μm pore size) seawater, concentrated (approximately 8 to 10 larvae per ml), and placed in polycarbonate petri dishes (17) mounted on plastic plates (13 by 13 cm; four dishes per plate). The larvae were allowed to settle overnight (approximately 150 per dish), and then half of the petri dishes containing developing colonies were treated with seawater containing 100 μg of gentamicin sulfate ml^{-1} for 7 days. The other half (controls) received unamended seawater. The seawater was changed and antibiotic solutions were prepared daily. A subset of the treated colonies was additionally treated with puromycin after 3 weeks of growth in an attempt to further eliminate "*E. sertula*." In preliminary studies, "*E. sertula*" had shown sensitivity to gentamicin, puromycin, and polymyxin B (6).

After the antibiotic treatment, the dishes were suspended vertically in tanks (25 to 40 liters) with aerated seawater maintained at 18 to 20°C and changed every 48 h with addition of non axenic phytoplankton (50% *Rhodomonas* sp. and equal portions of *Chlorella ellipsoidea*, *Nannochloris*, *Isochrysis*, and *Monochrysis* spp.) The colonies were maintained in fresh, nonsterile seawater which had been passed through a series of sand filters. Preliminary experiments had shown that it was not necessary to maintain the antibiotic-treated *B. neritina* in sterile seawater to prevent "*E. sertula*" from repopulating the colonies.

The cultures were examined weekly by eye and less frequently with a dissection microscope to monitor health and growth. Several dishes and at least 25 to 35 colonies from each treatment were examined. Periodically, healthy *B. neritina*

TABLE 1. Primers for amplification of KS domains

Primer designation	Sequence 5′-3′
KSD1F	ACR TGI GCR TTI GTI CC
KSD1R	ICA YGG IAC IGG IAC
KSa1F	ACG GAC AAG CGT CAT TAC
KSa1R	GCC AAG GCT TTA ATT CCG
KSa2F	GTT GTC TTT GCA GCA TCG CAT GTT ACC AC
KSa2R	CAC GCC CGC TAT CCC AGC ACC TAC C
T7KSa1F	TAA TAC GAC TCA CTA TAA CGG ACA AGC GTC ATT AC
T7KSa1R	TAA TAC GAC TCA CTA TAG CCA AGG CTT TAA TTC CG

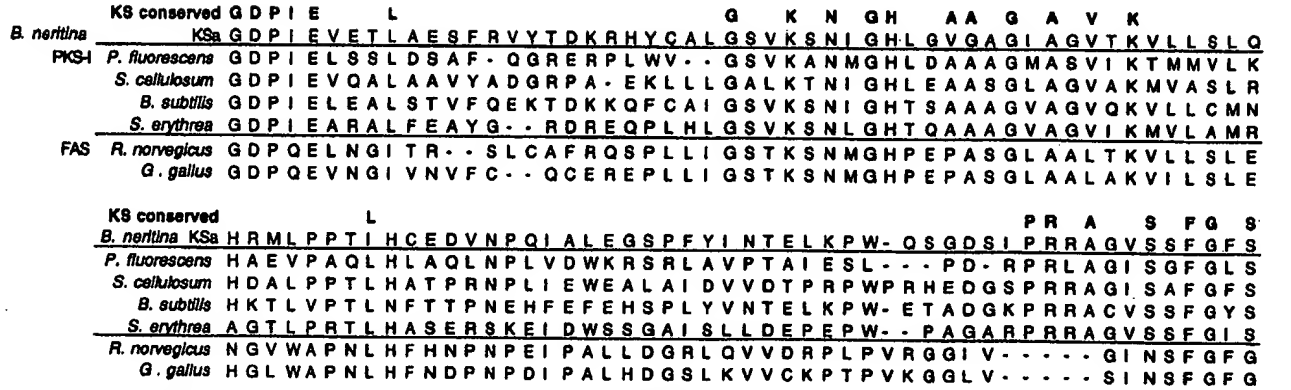

FIG. 2. Alignment of amino acid sequences of KSa from *B. neritina* with representative KS domains of PKS-I and FAS. Residues conserved among PKS-I are indicated above the *B. neritina* sequence.

colonies were removed from the dishes at the base using forceps for use in symbiont and bryostatin assays. Healthy colonies that contained live zooids were selected and pooled together for the assays.

Analysis of "*E. sertula*" levels, microbial community differences, and KSa signal. After growth in nonsterile seawater for 2 to 3 months, the *B. neritina* colonies had been colonized by bacteria from the seawater (growing on their surfaces and ingested). Denaturing gradient gel electrophoresis (DGGE) was used to compare the levels of both the "*E. sertula*" and the total bacterial communities between the control and treated colonies (11, 23). A portion of the small-subunit (SSU) rRNA gene was amplified using a conserved bacterial primer 1055f, and a universal primer with a GC clamp, 1392r-GC clamp (11). For amplification from *B. neritina*, 100 ng of DNA was used in 50-μl reaction mixtures with *Taq* polymerase and buffer (Boehringer Mannheim) as described by Haygood et al. (14), with 30 cycles of 94°C (1 min), 50°C (1 min), and 72°C (1 min). PCRs for each sample were run in triplicate and pooled prior to separation by DGGE. To make an "*E. sertula*" standard, symbiont SSU rRNA genes were amplified from *B. neritina* DNA using symbiont-specific primers as described by Haygood and Davidson (15). The "*E. sertula*" standard was then amplified with the primers described above from symbiont SSU rRNA gene amplification products and used to locate the "*E. sertula*" bands on gels.

Polyacrylamide gels (6.5%; 1:37 bis/acrylamide ratio) were prepared and run in TAE buffer (20 mM Tris base, 10 mM sodium acetate, and 0.5 mM EDTA, pH adjusted to 7.4) at 60°C with the CBS Scientific Inc. DGGE 2000 system at a running voltage of 200 V. An optimal gradient of 20 to 55% denaturant in the gel was used to separate the rRNA gene fragments. The DNA bands were visualized with 1× SYBR Gold (Molecular Probes, Eugene, Oreg.) on a UV transilluminator and were cut and extracted from the gel (11) for reamplification of DNA to be sequenced. Sequences from the suspected "*E. sertula*" band were compared with the "*E. sertula*" SSU rRNA sequence to confirm identity. The overall community was compared by examining the banding pattern from the control and treated samples.

Specific KSa primers (KSa1F and KSa1R) described above were used to amplify genes from equal amounts of DNA from the control and treated *B. neritina* to assess the level of this gene signal in treated colonies lacking most of their symbionts.

Bryostatin activity assays: tritiated phorbol dibutyrate displacement from protein kinase C (PKC) on rat brain liposomes. To measure bryostatin activity in small samples a method was adapted from the protocol described by Schaufelberger et al. (27) for detecting bryostatin activity in extracts (9, 10, 27). Bryostatins bind to PKC at the same site as phorbol ester, but with a much higher affinity under the assay conditions in which they irreversibly displace the phorbol from the binding site on PKC (9). This assay is not specific to bryostatins but indicates the presence of PKC binding and correlates well with bryostatin in *B. neritina* as determined by high-performance liquid chromatography (HPLC) fractionation of extracts from *B. neritina* (Mary Koleck [SAIC, Frederick, Md.] personal communication). The bryozoan *Scrupocellaria* sp. does not contain bryostatins and occasionally grows together with *B. neritina*. *Scrupocellaria* was used as a negative control for phorbol displacement activity in bryozoan tissue extracts.

To prepare *B. neritina* extracts, fresh or frozen (−80°C) samples were lyophilized and then homogenized in 95% ethanol for a final concentration of 12 μg ml^{-1} (dry weight to volume). Debris was pelleted by centrifugation, and the supernatants were used in assays. Each sample dose tested was run in triplicate to generate displacement curves.

A liposome suspension enriched in PKC was prepared by homogenization of one Sprague-Dawley rat brain in 30 ml of 10 mM Tris-HCL/buffer, pH 7.4, using a tissue tearor set at maximum (position 5, 1 min). The suspension was centrifuged at 16,000 × *g* for 10 min, and the supernatant was discarded. The process was repeated until the pellet was a buff color. The pellet was homogenized in 30 ml of 10 mM Tris buffer for 20 s on full speed with the tissue tearor, and then aliquots (1 to 2 ml) were stored at −80°C. Just before use, this suspension was thawed, homogenized, and centrifuged in a tabletop centrifuge at 14,000 rpm (20,800 × *g*) for 10 min. and then resuspended in the binding buffer (50 mM Tris-HCl [pH 7.4], bovine serum albumin [2 mg ml^{-1}]).

The crude extracts containing bryostatins together with tritiated phorbol dibutyrate (^{3}H-PDBu) were added to the suspension of rat brain liposomes rich in PKC. Reactions (1 ml total) were prepared in binding buffer with addition of extracts (between 2 and 15 μl), 10 μl of ^{3}H-PDBu (10 nM; 10 μCi), and 100 μl of the rat brain liposome suspension. Reactions with 10 μl of 20 μM unlabeled PDBu or no added sample were run as the positive and negative control solutions respectively. The negative control was run in duplicate for each assay and used as the total signal bound. Reactions were incubated for 1 h at 30°C and then passed through glass fiber filters (Whatman GF/B) on a vacuum manifold to capture liposomes, rinsed with 2 ml of 10 mM Tris-HCl buffer (pH 7.4), and placed in scintillation vials with 4 ml of scintillation cocktail. Vials were shaken for 30 min prior to reading on a scintillation counter. The reduction in tritium signal relative to the total tritium signal in the negative control was calculated and expressed as a percentage. The amount of crude extract required to displace 50% of the bound tritiated phorbol was the 50% inhibitory concentration (IC_{50}), which was the value compared between samples. Bryostatin 1 standard curves were generated to give the IC_{50} for bryostatin 1 and allowed an estimation of bryostatin levels (expressed as bryostatin 1 equivalents) in the samples.

Nucleotide sequence accession number. The nucleotide sequence of KSa from *B. neritina* (Fig. 2) has been assigned GenBank accession no. AF283572.

RESULTS

Amplification, cloning, and sequencing of KS domain DNA from *B. neritina*. Degenerate primers (KSD1F and KSD1R) were used in a step-down PCR protocol to amplify a 300-bp fragment from *B. neritina* DNA. Since bacterial PKS-I enzymes are modular, PCR products represent a pool of KS fragments from different modules and different bacteria present in the sample. Two clone libraries were prepared, one from adult *B. neritina* DNA and one from larval DNA. Twenty-seven clones were sequenced, and nine unique sequences were identified, designated KSa to KSi. Four unique sequences, KSa (13 clones), KSb (1 clone), KSe (3 clones), and KSh (1 clone), were obtained from the larval DNA library, and six sequences were

a. Torrey Pines, 9-12 m, 6/97
b. Scripps Pier, 0-7.6 m, 6/97
c. Palos Verdes, 16.5 m, early '97
d. Palos Verdes, Larvae, early '97
e. Catalina Island, Blue caverns Pt., 9-12 m, 8/97
f. Catalina Island, Harbor Dock, surface, 8/97
g. Catalina Island, Larvae, 8/97
h. Newport, CI, Boat hull, 0-2 m, 8/97
i. Humboldt Bay Dock, surface, 9/97
j. Tomales Bay, 0-1.5 m, 9/97
k. Patrick's Point (outside Humboldt Bay), 3 m, 9/97
l. Bodega Bay Dock, surface, 9/97
m. North Carolina (Atlantic), 2 m, fall '97
n. Scrupocellaria sp., Scripps Pier, '97
o. Unidentified white bryozoan,Scripps Pier, 3/98

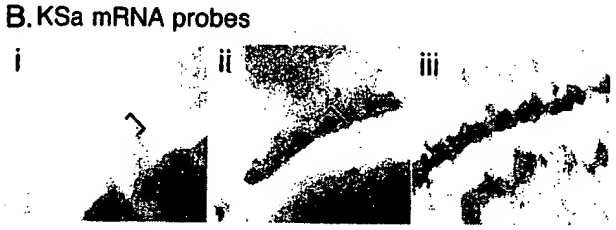

FIG. 3. KSa detection in *B. neritina* and in bacteria of the larvae. (A) Amplification of a KSa gene fragment from *B. neritina* DNA using specific primers. Lanes A to L show amplification of the KSa gene fragment in a variety of *B. neritina* samples, including both adults and larvae, collected throughout the year from different depths and habitats in California, and lane M shows one sample from North Carolina. Lanes N, *Scrupocellaria* sp., and O, unidentified white bryozoan, show lack of KSa amplification from two other species of bryozoa collected with *B. neritina* from the Scripps Pier. (B) In situ hybridization of KSa RNA probe in whole *B. neritina* larvae. The pallial sinus containing bacterial cells is indicated by a bracket. (i) sense KSa probe (negative control); (ii) antisense KSa probe; (iii) 16S rRNA probe showing location of "*E. sertula*" bacterial cells in the pallial sinus.

obtained from the adult library, KSb, KSc, KSd (two clones each), KSf, KSg, and KSi (one clone each). Only the KSb sequence was found in both libraries. Cloned DNA sequences were analyzed by using the BLAST (Basic Local Alignment Search Tool) server of the National Center for Biotechnology Information, accessed over the Internet (1). All of these sequences have the expected conserved signature regions for KS and show highest similarity in BLAST searches to bacterial PKS-I (Fig. 2). The most abundant sequence, KSa, was present in 13 out of 19 clones from the larval library (GenBank accession number AF283572) and was targeted for further study. An alignment of 85 amino acids of the clones, unalignable regions excluded, revealed identities between 32 and 58% relative to KSa.

Survey of *B. neritina* for KSa content and in situ hybridization. The KSa gene fragment was amplified by PCR from 10 samples of DNA from adult *B. neritina* colonies found in a wide range of locations in California from Humbolt Bay to San Diego, and included Catalina Island. Various habitat types and depths were sampled (boat docks and piers, 1 to 2 m; submerged rocks, 16 to 25 m). In addition, *B. neritina* from North Carolina was included. All of these samples had been shown previously to contain bryostatins (7). All samples produced ample product when KSa was amplified with the KSa-specific primers (Fig. 3A, lanes A to M). Two samples of other bryozoans collected with *B. neritina* did not yield products (Fig. 3A, lanes N and O). In similar experiments with the other eight KS sequences cloned, KSb to KSi, bands were only observed in a few samples, suggesting that these sequences originated in environmental bacteria.

An antisense KSa RNA probe bound to the bacterial cells in the pallial sinus of the larvae, indicating that the "*E. sertula*" symbiont expresses the KSa message (Fig. 3B, panel ii). This probe did not bind to *E. coli* (not shown). The sense probe did not bind to bacterial cells in the larvae (Fig. 3B, panel i).

Growth of *B. neritina* colonies after antibiotic treatments. After antibiotic treatments, the growth of the treated and that of the control *B. neritina* colonies were similar, showing no apparent detrimental effects from either the gentamicin treatments (up to 150 $\mu g\ ml^{-1}$) or the reduction of symbiont population. The zooids, lophophores, and growth form of the colonies appeared normal relative to the control colonies. In all cases the range of colony sizes increased as growth progressed. The range of sizes and average size were similar between the control and the gentamicin-treated colonies (Fig. 4A) within a settlement. Between settlements (1996, 1997, and 1998) there was some difference in growth rates and final colony sizes reached; however, in every case the matched control and treated batches within a settlement had similar growth rates and final colony sizes. The differences between the settlements were likely due to variability of the larvae and environmental

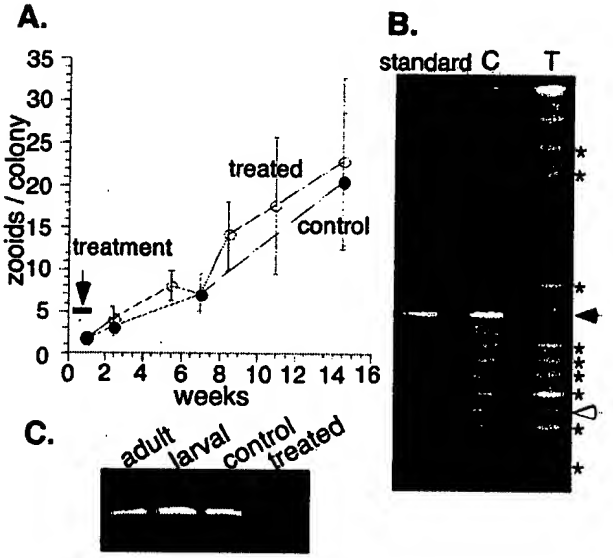

FIG. 4. Growth, bacterial content and KSa gene detection in control and treated *B. neritina*. (A) Growth of colonies in number of zooids per colony. Thickbar; period of antibiotic treatment error bars, standard deviation. (B) Denaturing gradient gel showing bacterial SSU rRNA fragment pattern from treated and control colonies. Standard, purified "*E. sertula*"; solid arrow, "*E. sertula*"; *, common fragments; open arrow, band appearing in control diminished in treated *B. neritina* DNA. (C) PKS KSa gene fragments amplified from DNA of *B. neritina* wild adults, larvae, and laboratory-cultured control and antibiotic-treated colonies.

conditions, since both the control and treated *B. neritina* behaved in a similar manner.

"*E. sertula*" DNA levels and microbial community analysis of treated and control colonies. After an initial treatment with 100 μg of gentamicin sulfate ml^{-1} for 7 days, colonies were allowed to grow without further antibiotic treatments, except for a subset of the 1998 settlement which underwent a second antibiotic treatment. Symbiont levels were analyzed at 1 month and 2 to 3 months. Several replicate analyses of treated and control samples were conducted with consistent results. A representative DGGE gel is shown in Fig. 4B.

DGGE of 16S rRNA amplification products showed that the banding patterns were similar between the control and treated colony DNA amplifications, except for the "*E. sertula*" band (Fig. 4B). The "*E. sertula*" signal was by far the strongest in the control colony DNA, suggesting that this was the most-abundant bacterial species associated with *B. neritina*. The "*E. sertula*" signal was greatly reduced in the treated-colony DNA. In addition, a few bands appeared more strongly in the treated-colony DNA lane than they appear in the control lane. All of the amplification products that appeared in the controls also appeared in the treated-colony lane, with the exception of the "*E. sertula*" band and one additional faint band that was also reduced. This second amplification product was not consistently detected in wild specimens (6). This pattern indicated that the only major bacterial species reduced by the gentamicin treatments was the symbiont "*E. sertula*." The decrease in the "*E. sertula*" population was confirmed by comparing the PCR products amplified with "*E. sertula*" -specific primers from a dilution series of quantified treated-colony DNA with dilutions of quantified control colony DNA. These comparisons showed a 50- to 100-fold reduction of "*E. sertula*" DNA in treated colonies relative to controls (250 ng of treated-colony DNA yielded product comparable to between 2.5 and 5 ng of control-colony DNA [6]).

Detection of polyketide synthase gene KSa. PCR gene amplification using KSa-specific primers from wild adult, larvae, and cultured control *B. neritina* colony DNA all yielded strong signals for the PKS KSa gene fragment. Amplification from the treated-colony DNA resulted in only a small amount of product (Fig. 4C), indicating a substantially reduced level of KSa present in treated colonies.

Bryostatin activity assays. Extracts of *B. neritina* colonies from three gentamicin-treated and control settlements were tested for bryostatin activity using the tritiated phorbol displacement assay. In addition, colonies of *Scrupocellaria* sp. that had fortuitously settled in the dishes along with the *B. neritina* larvae were assayed for activity. The *Scrupocellaria* sp. extracts showed no phorbol displacement activity (Fig. 5A).

A representative graph of the percentage of tritiated phorbol displacement plotted as a function of the amount of sample added is shown in Fig. 5A. The IC_{50}s calculated from the binding curves were expressed as IC_{50}^{-1} (activity per microgram of sample). The IC_{50}s indicated that reduction of the "*E. sertula*" population was consistently correlated with reduced bryostatin levels (22 to 60% less bryostatin activity by weight). The least difference in bryostatin activity between control and treated *B. neritina* was observed when the colonies had been growing quickly and the level of bryostatins was low in the control colonies (1998 settlement). It is possible that the levels of bryostatin production could not keep up with the rate of bryozoan colony growth. The greatest

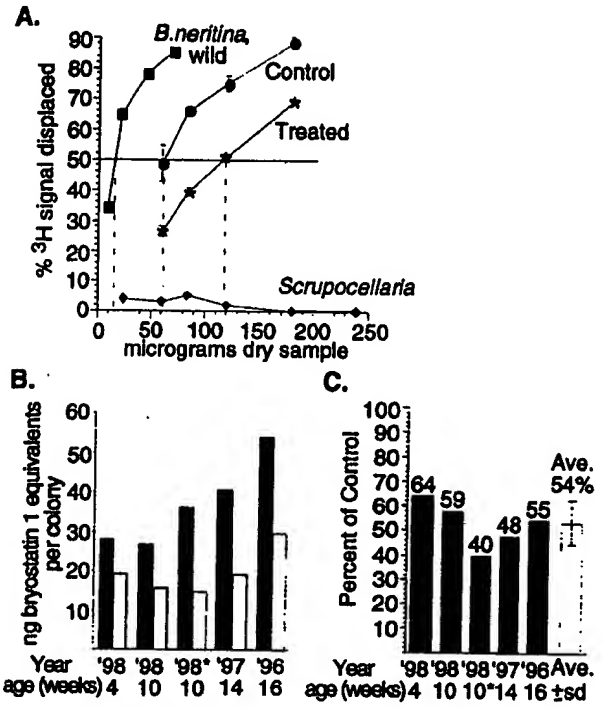

FIG. 5. Bryostatin activity. (A) Representative ^{3}H-PDBu displacement curve used to determine relative bryostatin activity levels in *B. neritina*. Standard deviations (error bars) are shown for each point; however, for most points the variation is only as wide as or smaller than the point symbol. Dashed lines dropped to the *x* axis indicate IC_{50}s (dose giving 50% displacement). *Scrupocellaria* sp. extracts were tested as a negative control. (B) Bryostatin 1 equivalents in laboratory-grown *B. neritina* control (solid bars) and treated (open bars) colonies from settlements from 1996 to 1998 (the bryostatin 1 PDBu displacement curve gave an IC_{50} of 13 ng). (C) The percentage of bryostatin activity remaining in treated colonies relative to controls for all treatments. The bryostatin activity levels in controls represent the 100% values. For all values, variation was less than 1.5 percentage points. Open bar, the average bryostatin level remaining in treated colonies for all settlements and ages, with the standard deviation (error bar) shown.

differences in bryostatin activity were observed when growth had slowed, as in the 1996 end point measurements and the subset of 1998 settlement treated a second time with antibiotics. A likely explanation is that the bryostatins were accumulating in the slowly growing colonies but accumulated at a faster rate in the controls. Since the measurements are by weight, if growth slows and the production of bryostatins remains the same, then bryostatins will accumulate faster in colonies with a slow growth rate. The highest bryostatin levels by weight and the greatest difference between the control and treated-colony bryostatin activities occurred with the 1998 double antibiotic treatment, in which the growth was stunted but the colonies appeared healthy and were feeding. In this treatment the growth of both the control and treated colonies slowed due to the removal from the main tank for 1 week. The sizes of both control and treated colonies were similar.

The IC_{50}s were transformed into bryostatin 1 equivalents by comparing the IC_{50}s of the colony extracts to that of bryostatin 1. Since the PDBu assay measures binding to PKC by all bryostatins present in the sample, this number is referred to as the bryostatin 1 equivalent and does not represent the actual

amount of bryostatin 1. The amounts by weight were converted into amount per zooid from determined dried zooid weights. The total amount of bryostatin per colony was estimated based on colony size (zooid number). We found that the average of the three treatments showed that 54% of the bryostatin activity remained in treated colonies when comparing bryostatin 1 equivalents per colony.

DISCUSSION

The biosynthetic capability for production of bryostatins must reside in the bryozoan *B. neritina*, its specific bacterial symbiont "*E. sertula*," another associated organism, or a combination of these organisms. In this article, we describe evidence that "*E. sertula*" has the genetic ability to produce complex PKS-I and that the symbiont is required for bryostatin synthesis, suggesting that "*E. sertula*" is the biosynthetic source of bryostatins.

Gene fragments from the KS domain, which is present in all modules of bacterial PKS-I genes, were amplified, cloned, and sequenced from both larval and adult *B. neritina* DNA. The most abundant sequence, KSa, was found in 13 out of 19 larval clones (Fig. 2). The abundance of KSa sequence among the clones obtained from DNA of the larvae suggest "*E. sertula*" as the source of this sequence. Larvae should have primarily "*E. sertula*" bacterial DNA since the larvae have a monoculture of the symbiont in the pallial sinus (15) and do not have a gut. PCR using primers (Table 1) specific to the KSa clone amplified this gene fragment from all adult and larval *B. neritina* samples tested but did not result in amplification from other bryozoans growing in proximity to *B. neritina* (Fig. 3A). This survey included samples from the Atlantic; Northern and Southern California; boat docks; submerged rocks; and different seasons, temperatures, and depths. The presence of the Ksa gene fragment in all *B. neritina* samples showed that the sequence is specific to and always present in *B. neritina*. If KSa originated from a contaminant found in the environment we would expect the presence of this gene fragment to have varied in different samples, as was the case with the other eight KS sequences cloned, and to have appeared in other bryozoans living in the same milieu as *B. neritina*.

An RNA antisense probe synthesized from a KSa clone was used for in situ hybridization to label KSa message within the larvae. This probe bound to bacterial cells within the pallial sinus, indicating that the mRNA for KSa was expressed in the "*E. sertula*" cells of the pallial sinus of *B. neritina* (Fig. 3B). The negative control sense probe did not bind. The KSa probe did not bind to closely related bacteria, indicating that the probe was not binding to an FAS gene but was specific to "*E. sertula*" (not shown). Although we do not know whether these PKS genes are involved in the bryostatin synthesis pathway, the presence of these genes in "*E. sertula*" shows that they have the potential for synthesis of complex polyketides such as bryostatin.

To test the hypothesis that "*E. sertula*" is involved in bryostatin production in the host bryozoan, and to confirm by a third method that KSa originated from "*E. sertula*," we attempted to eliminate the symbiont from *B. neritina* reared in the laboratory and assayed for bryostatin and KSa gene levels. To date it has not been possible to eliminate the symbiont from the host entirely. However, the levels of "*E. sertula*" were substantially reduced (an estimated 5% remaining) in gentamicin-treated *B. neritina* as shown by DGGE comparative analysis of total SSU rRNA genes amplified from control and treated colony DNA (Fig. 4B). This result was confirmed by specific amplification of "*E. sertula*" SSU rRNA genes from a dilution series of control and treated-colony DNA (6). Although these methods are not strictly quantitative, the results from several amplifications were consistent and confirmed that the "*E. sertula*" was substantially reduced by antibiotic treatment.

The bacteria that colonize *B. neritina* after settlement are not likely specific symbionts but environmental bacteria that colonize surfaces and pass through the gut tract of the bryozoan. DGGE analysis of SSU rRNA genes confirmed that the total microbial community, excluding "*E. sertula*," was similar in treated and control colonies (Fig. 4B). "*E. sertula*" yielded the dominant signal in control colonies and was greatly reduced in the treated-colony bacterial community SSU rRNA profiles. SSU rRNA genes amplified from the control colonies were also present in the treated colonies, indicating that similar microbial species were present in both. Although one additional minor product was detected in the controls that was reduced in the treated colonies, this product has not consistently appeared in DGGE analyses of wild populations (6). This sporadic appearance indicated that this bacterium was not a candidate for the bryostatin producer. To date DGGE analyses of wild *B. neritina* have shown that only "*E. sertula*" was consistently found associated with *B. neritina* (6), although it is possible that there is another symbiont present that is not detectable by PCR of the SSU rRNA genes. PCR amplification of the KSa gene using specific primers yielded a strong signal from control *B. neritina* colonies but produced a weak signal from the treated colonies with reduced "*E. sertula*" (Fig. 4C). In light of the DGGE evidence, the reduction of KSa in the treated *B. neritina* is due to the reduction of "*E. sertula*" numbers.

PDBu displacement assays indicated an average reduction of 46% in bryostatin activity from treated colonies relative to controls (Fig. 5C). This reduction, however, was not proportional to the loss of bacterial symbiont (46% loss in activity compared to 95% estimated loss of "*E. sertula*"). This finding may seem to contradict the hypothesis that the bacterial symbionts produce bryostatins, but there are reasons not to expect a strict linear relationship between the two. The synthesis of bryostatins may be under the control of the symbiont or the host and may be up-regulated when bacterial numbers are reduced. An alternative explanation is that there are other compounds in the bryozoan or from another undetected symbiont that bind PKC. The data currently available do not support either of these alternative explanations. Extracts of other bryozoans have not shown activity in the PKC binding assay. HPLC peaks from *B. neritina* extracts that test positive in this assay have been identified as bryostatins. *B. neritina* compounds have been investigated for their PKC binding ability and bioactivity for the past 30 years, and no other compounds besides bryostatins have been found. Considering these factors, it is unlikely that there are other compounds that react in the PKC binding assay to the level we observed (54% of the activity). The possibility of an undetected symbiont has been investigated by bacterial population surveys of *B. neritina* from different locations, and data to date do not support a second unidentified bacterial symbiont. In addition, the larvae contain one transmitted symbiotic species, "*E. sertula*," and contain bryostatins.

Since "*E. sertula*" was the only significantly reduced bacterial species after the antibiotic treatments and since the growth and morphology of treated colonies showed no detrimental effects, differences in KSa gene signal and bryostatin activity can be attributed to the reduction of "*E. sertula*."

In summary, an abundant KS sequence, KSa, was (i) shown by PCR to be present in all populations of *B. neritina* available, (ii) shown by mRNA hybridization to be expressed in the "*E. sertula*" cells in the pallial sinus of the larvae, and (iii) shown to be reduced in concert with reduction of "*E. sertula*" levels. These data indicate that the KSa sequence originates in the bacterial symbiont "*E. sertula*." In addition, the observation that reduction of "*E. sertula*" levels reduced bryostatin production suggests that "*E. sertula*" is the likely biosynthetic source of the bryopyran ring and possibly the entire bryostatin molecule. This finding has important implications for understanding the function of the symbiont in this association between a bryozoan and a bacterium. Although the benefit of bryostatins to the bryozoan remains untested, bryostatins likely serve as a form of chemical defense that is provided by the bacterial symbionts. These findings also illustrate how molecular methods may be used to solve some of the problems associated with advancing marine natural products from discovery to use as therapies. If symbiotic bacteria or their genes can be harnessed for production of bryostatins or novel analogues thereof, lack of supply may no longer be a barrier for the development and use of this important family of compounds.

ACKNOWLEDGMENTS

CalBioMarine Technologies, Inc. (Carlsbad, Calif.) contributed larvae and expertise regarding the culturing of *B. neritina*. Carolyn Sheehan provided technical assistance. Thanks also go to Ron McConnaughey, who assisted with collections of *B. neritina* adults and helped set up aquaria for growth experiments.

This work was supported by a California Sea Grant College grant (R/MP-61) and the National Sea Grant Technology Program (R/MP-84A). S. W. Allen was supported on a Small Business Innovation Research grant (R44 CA58158-02A3) from the National Cancer Institute, awarded to CalBioMarine Technologies, Inc. G. E. Lim was supported by a Howard Hughes Medical Institute Fellowship.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.
2. **Anthoni, U., P. H. Nielsen, M. Perieira, and C. Christophersen.** 1990. Comp. Biochem. Physiol. 96B:431–437.
3. **Baldwin, N. G., C. D. Rice, T. M. Tuttle, H. D. Bear, J. L Hirsch, and R. E. Merchant.** 1997. Ex vivo expansion of tumor-draining lymph node cells using compounds which activate intracellular signal transduction. I. Characterization and in vivo anti-tumor activity of glioma-sensitized lymphocytes. J. Neurooncol. 32:19–28.
4. **Basu, A.** 1998. The involvement of novel protein kinase C isozymes in influencing sensitivity of breast cancer MCF-7 cells to tumor necrosis factor-alpha. Mol. Pharmacol. 53: 105–111.
5. **Correale, P., M. Caraglia, A. Fabbrocini, R. Guarrasi, S. Pepe, V. Patella, G. Marone, A. Pinto, A. R. Bianco, and P. Tagliaferri.** 1995. Bryostatin 1 enhances lymphokine activated killer sensitivity and modulates the beta 1 integrin profile of cultured human tumor cells. Anticancer Drugs 6:285–90.
6. **Davidson, S. K.** 1999. The biology of the bryostatins in the marine bryozoan *Bugula neritina*. Dissertation. University of California, San Diego, La Jolla, Calif.
7. **Davidson, S. K., and M. G. Haygood.** 1999. Identification of sibling species of the bryozoan *Bugula neritina* that produce different anticancer bryostatins and harbor distinct strains of the bacterial symbiont "*Candidatus* Endobugula sertula." Biol. Bull. 196:273–280.
8. **DeLong, E. F., L. T. Taylor, T. L. Marsh, and C. M. Preston.** 1999. Visualization and enumeration of marine planktonic archaea and bacteria by using polyribonucleotide probes and fluorescent in situ hybridization. Appl. Environ. Microbiol. 65:5554–5563.
9. **DeVries, D. J., C. L. Herald, G. R. Pettit, and P. M. Blumberg.** 1988. Demonstration of sub-nanomolar affinity of bryostatin 1 for the phorbol ester receptor in rat brain. Biochem. Pharmacol. 37:4068–4073.
10. **DeVries, D. J., K. S. Rao, and R. H. Willis.** 1997. Application of a radioreceptor assay to the screening and characterisation of compounds from marine organisms with activity at the phorbol ester binding site of protein kinase C. Toxicon 35:347–354.
11. **Ferris, M. J., G. Muyzer, and D. M. Ward.** 1996. Denaturing gradient gel electrophoresis profiles of 16S rRNA-defined populations inhabiting a hot spring microbial mat community. Appl. Environ. Microbiol. 62:340–346.
12. **Fleming, M. D., H. D. Bear, K. Lipshy, P. J. Kostuchenko, D. Portocarero, A. W. McFadden, and S. K. Barrett.** 1995. Adoptive transfer of bryostatin-activated tumor-sensitized lymphocytes prevents or destroys tumor metastases without expansion in vitro. J. Immunother. Emphasis Tumor Immunol. 18:147–55.
13. **Grant, S., R. Traylor, G. R. Pettit, and P. S. Lin.** 1994. The macrocyclic lactone protein kinase C activator, bryostatin 1, either alone, or in conjunction with recombinant murine granulocyte-macrophage colony-stimulating factor, protects Balb/c and C3H/HeN mice from the lethal *in vivo* effects of ionizing radiation. Blood 83:663–667.
14. **Haygood, M., D. Distel, and P. Herring.** 1992. Polymerase chain reaction and 16S rRNA gene sequences from the luminous bacterial symbionts of two deep-sea anglerfishes. J. Mar. Biol. Assoc. U. K. 71:149–159.
15. **Haygood, M. G., and S. K. Davidson.** 1997. Small subunit ribosomal RNA genes and *in situ* hybridization of the bacterial symbionts in the larvae of the bryozoan *Bugula neritina* and proposal of "*Candidatus* Endobugula sertula." Appl. Env. Microbiol. 63:4612–4616.
16. **Johnson, M. D., J. A. Torri, M. E. Lippman, and R. B. Dickson.** 1999. Regulation of motility and protease expression in PKC-mediated induction of MCF-7 breast cancer cell invasiveness. Exp. Cell Res. 247:105–113.
17. **Keough, M. J.** 1989. Dispersal of the bryozoan *Bugula neritina* and effects of adults on newly metamorphosed juveniles. Mar. Ecol. Prog. Ser. 57:163–171.
18. **Kraft, A. S., J. B. Smith, and R. L. Berkow.** 1986. Bryostatin, an activator of the calcium phospholipid-dependent protein kinase, blocks phorbol ester-induced differentiation of human promyelocytic leukemia cells HL-60. Proc. Natl. Acad. Sci. USA 83:1334–1338.
19. **Kraft, A. S., S. Woodley, G. R. Pettit, F. Gao, J. C. Coll, and F. Wagner.** 1996. Comparison of the antitumor activity of bryostatins 1, 5, and 8. Cancer Chemother. Pharmacol. 37:271–278.
20. **Lind, D. S., T. M. Tuttle, K. P. Bethke, J. L. Frank, C. W. McCrady, and H. D. Bear.** 1993. Expansion and tumor specific cytokine secretion of bryostatin-activated T-cells from cryopreserved axillary lymph nodes of breast cancer patients. Surg. Oncol. 2:273–282.
21. **Lipshy, K. A., P. J. Kostuchenko, G. G. Hamad, C. E. Bland, S. K. Barrett, and H. D. Bear.** 1997. Sensitizing T-lymphocytes for adoptive immunotherapy by vaccination with wild-type or cytokine gene-transduced melanoma. Ann. Surg. Oncol. 4:334–341.
22. **Maloy, S. R., V. J. Stewart, and R. K. Taylor.** 1996. Genetic analysis of pathogenic bacteria. Cold Spring Harbor Laboratory Press, Plainview, N.Y.
23. **Muyzer, G., E. C. D. Waal, and A. G. Uitterlinden.** 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. 59:695–700.
24. **Pettit, G. R., C. L. Herald, D. L. Doubek, and D. L. Herald.** 1982. Isolation and structure of bryostatin 1. J. Am. Chem. Soc. 104:6846–6848.
25. **Pettit, G. R.** 1991. The bryostatins. Prog. Chem. Org. Nat. Prod. 57:153–195.
26. **Pluda, J. M., B. D. Cheson, and P. H. Phillips.** 1996. Clinical trials referral resource. Clinical trials using bryostatin-1. Oncology (Huntington) 10:740–742.
27. **Schaufelberger, D. E., A. B. Alvarado, P. Andrews, and J. A. Beutler.** 1990. Detection and quantitation of bryostatin 1 and 2 in *Bugula neritina* by combined high-performance liquid chromatography and 3H-phorbol dibutyrate displacement. J. Liq. Chromatography 13:583–598.
28. **Scheid, C., J. Prendiville, G. Jayson, D. Crowther, B. Fox, G. R. Pettit, and P. L. Stern.** 1994. Immunomodulation in patients receiving intravenous bryostatin 1 in a phase I clinical study: comparison with effects of bryostatin 1 on lymphocyte function in vitro. Cancer Immunol. Immunother. 39:223–230.
29. **Shure, M., S. Wessler, and N. Fedoroff.** 1983. Molecular identification and isolation of the waxy locus in maize. Cell 35:225–233.
30. **Smith, J. B., L. Smith, and G. R. Pettit.** 1985. Bryostatins: potent, new mitogens that mimic phorbol ester tumor promoters. Biochem. Biophys. Res. Commun. 132:939–945.
31. **Steube, K. G., and H. G. Drexler.** 1993. Differentiation and growth modulation of myeloid leukemia cells by the protein kinase C activating agent bryostatin-1. Leuk. Lymphoma 9:141–148.
32. **Sung, S. J., P. S. Lin, R. Schmidt-Ullrich, C. E. Hall, J. A. Walters, C. McCrady, and S. Grant.** 1994. Effects of the protein kinase C stimulant bryostatin 1 on the proliferation and colony formation of irradiated human T-lymphocytes. Int. J. Radiat. Biol. 66:775–783.
33. **Taylor, L. S., G. W. Cox, G. Melillo, M. C. Bosco, and L. Espinoza-Delgado.** 1997. Bryostatin-1 and IFN-gamma synergize for the expression of the inducible nitric oxide synthase gene and for nitric oxide production in murine macrophages. Cancer Res. 57:2468–73.
34. **Woollacott, R. M.** 1981. Association of bacteria with bryozoans larvae. Mar. Biol. 65:155–158.

Modular Organization of Genes Required for Complex Polyketide Biosynthesis

STEFANO DONADIO, MICHAEL J. STAVER, JAMES B. MCALPINE, SUSAN J. SWANSON, LEONARD KATZ

In *Saccharopolyspora erythraea*, the genes that govern synthesis of the polyketide portion of the macrolide antibiotic erythromycin are organized in six repeated units that encode fatty acid synthase (FAS)–like activities. Each repeated unit is designated a module, and two modules are contained in a single open reading frame. A model for the synthesis of this complex polyketide is proposed, where each module encodes a functional synthase unit and each synthase unit participates specifically in one of the six FAS-like elongation steps required for formation of the polyketide. In addition, genetic organization and biochemical order of events appear to be colinear. Evidence for the model is provided by construction of a selected mutant and by isolation of a polyketide of predicted structure.

POLYKETIDES ARE A LARGE AND DIVERSE CLASS OF NATURAL products that includes antibiotics, pigments, and immunosuppressants and have applications in medicine, agriculture, and industry (for an example, see Fig. 1). Biosynthesis of polyketides is believed to occur by a series of condensations of carbon units in a manner similar to that of long chain fatty acids (LCFA) (*1*). The LCFAs are formed by fatty acid synthase (FAS) through a process whereby a starter unit (commonly acetate) is condensed to the extender unit (malonate). The resulting β-keto group is then fully processed (reduced), and the cycle resumes with the condensation of a new extender unit (*2*). Most polyketides, however, contain structural complexities that can be accounted for by the use of different extender units at various steps and by variations in the extent of processing of the β-carbon (β-ketoreduction, dehydration, enoylreduction). Although this complexity exists, in living organisms there is believed to be a polyketide synthase (PKS) that produces only one or a few related molecular structures (*3*). Thus, an understanding of the biosynthesis of complex polyketides must include a description of the mechanism by which the PKS both selects the correct substrate and decides the fate of the β carbon at each step.

The polyketide portion of macrolide antibiotics is synthesized through the condensation of short chain carbon units; for example, seven propionates in the case of erythromycin (*4*). β-Hydroxy-acyl thioesters, which mimic hypothetical intermediates in the synthesis but not the corresponding β-keto derivatives, are incorporated in vivo into the macrolide rings of erythromycin and tylosin, providing support for a FAS-like origin of these molecules (*5*). Consistent with this idea is the detection of branched-chain fatty acids in fermentation broths of tylosin and mycinamicin producers (*6*). A FAS-like mechanism for the synthesis of the erythromycin aglycone 6-deoxyerythronolide B (6dEB) requires that (i) six methylmalonyl–coenzyme A (mmCoA) units, three of each enantiomer, are successively condensed to a propionyl-CoA starter unit; (ii) that β-ketoreduction occurs after each condensation step except step three so that a keto group is left at C-9; and (iii) that dehydration and enoyl reduction take place only after the fourth condensation to introduce a methylene at C-7 (Fig. 1). Consequently, a full set of FAS activities (*2*) is required for 6dEB synthesis: acyltransferase (AT), β-ketoacyl carrier protein synthase (KS), and acyl carrier protein (ACP) for chain elongation; β-ketoreductase (KR), dehydratase (DH), and enoyl reductase (ER) for processing of the β carbon; and thioesterase (TE) for release and lactonization of the full-length chain. In addition, the hypothetical 6dEB PKS must also be programmed at each step to select the correct enantiomeric extender unit and to process the β carbon to the appropriate degree.

Six modules with FAS-like domains in 6dEB synthesis. The segment of the chromosome required for the formation of 6dEB in *S. erythraea* has been designated *eryA*. A 5-kb DNA fragment of *eryA* had been identified by its ability to restore erythromycin production when introduced into a mutant blocked in the synthesis of 6dEB. Hybridization of chromosomal DNA with this segment

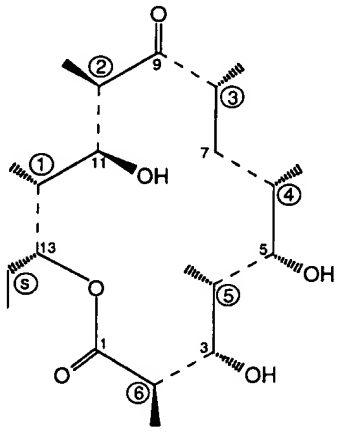

Fig. 1. Structure of 6-deoxyerythronolide B. The dotted lines represent C-C bonds formed during synthesis steps 1 through 6 by condensation between the starter unit propionyl-CoA (encircled S) and the extender units methylmalonyl-CoA encircled 1 to 6. The stereochemistry of extender units 1, 3, and 4 should be compared to that of units 2, 5, and 6. Lactonization of the acyl chain between C-1 and C-13 results in the formation of 6-deoxyerythronolide B. Hydroxylation of the lactone ring at C-6, followed in order by mycarose and desosamine attachment at hydroxyls at C-3 and C-5, respectively, results in the formation of the first bioactive compound, erythromycin D (*4*).

S. Donadio, M. J. Staver, and L. Katz are in the Department of Corporate Molecular Biology, Abbott Laboratories, Abbott Park, IL 60064. J. B. McAlpine and S. J. Swanson are in the Bioactive Microbial Metabolite Project, Abbott Laboratories, Abbott Park, IL 60064.

has indicated the existence of several closely linked homologous loci (7). Gene disruption experiments (7, 8) showed that *eryA* encompasses about 10 kb of DNA and identified the end of *eryA* proximal to *ermE*, the erythromycin resistance gene (9). Similar experiments were performed in our study to establish the other boundary for *eryA*. Contiguous or overlapping DNA segments that spanned a region of 25 kb were subcloned into the poorly replicating *S. erythraea* vector pWHM3 (10), and the resulting plasmids were integrated into the wild-type chromosome by single reciprocal recombination (11). Except for the *ermE*-distal fragment (Fig. 2A), all integrants exhibited an EryA phenotype (12). Because a mutant phenotype is expected when the cloned fragment is internal to a transcription unit, these results indicate that *eryA* extends for 32 to 35 kb and is likely to consist of one or possibly two transcripts. No integration events were obtained in a 1.6-kb segment of *eryA*, but this segment is fully internal to *orf1*; see Fig. 2C. Nucleotide sequencing (13) of the 35-kb segment established that *eryA* consists of six repeated units. Comparison of the 35-kb segment of sequenced DNA with a 4.5-kb subsegment of *eryA* indicates that the sizes of each of the six repeated units is 4.3 to 6.5 kb and that the sequences have a similarity of 64 percent or higher (Fig. 2B).

Three large open reading frames (*orf's*) are present in *eryA*, each consisting of two repeated units (Fig. 2C). The ORF organization of *eryA* supports the possible existence of large transcripts, which is indicated by gene disruption experiments. The putative start for *orf1* is within the 3-kb segment (full bar in Fig. 2A) where one of the *eryA* boundaries was mapped by gene disruption. A segment of 1.44 kb apparently not involved in 6dEB synthesis separates *orf1* and *orf2*, whereas *orf2* and *orf3* are contiguous; *orf3* corresponds to ORFA described independently (14). The deduced amino acid sequences of the three Orf's were compared to FAS and PKS systems, and putative FAS-like domains in *eryA* are depicted at their approximate map positions (Fig. 2C). A total of seven ACP's, six KS's, eight AT's, six KR's, one DH, and one ER have been identified and are organized into the six repeated units (Fig. 2B). The segment of *eryA* encoding a repeated unit is designated a "module" (Fig. 2C). Starting from the 5′ end of *eryA*, the first module encodes putative activities in the order AT, ACP, KS, AT, KR, and ACP. The second, third, and fifth modules consist of KS, AT, KR, and ACP. The fourth contains KS, AT, DH, ER, KR, and ACP, and the sixth, KS, AT, KR, ACP, and AT. In summary, *eryA* consists of six modules that encode FAS-like activities, and each Orf contains two modules.

Each of the putative domains in the six modules contains the active site motifs reported to be conserved in FAS and PKS systems (15). Specifically, the acyltransferase GHSxG (15; x = any amino acid) motif embedded in a hydrophobic region (where S is the serine involved in the formation of the acyl-enzyme intermediate) is present in each of the eight proposed *eryA* AT's. Similarly, the highly conserved sequence GPxxxxTACSS around the cysteine residue that participates in thioester formation can be detected in the six putative KS domains of *eryA*. The pantotheine-binding serine, present in most ACP's in the LGxDSLxxVE motif, is found in the seven *eryA* ACP's identified, and the "fingerprint" region GxxGxx-AxxA of NADPH (reduced form of nicotinamide adenine dinucleotide phosphate)-dependent reductases (16) is maintained in similar surroundings in the *eryA* KR domains. The functional domains for KR, DH and ER were identified as described below.

Our experiments extend to the three *eryA* Orf's recent findings (14), which showed that the predicted polypeptide designated Orf3 here resembles the type I eukaryotic FAS systems, wherein the functions required for fatty acid synthesis reside within a single polypeptide. This contrasts with the type II–like organization (monofunctional polypeptides) of PKS's involved in the synthesis of aromatic polyketides in *Streptomyces* (17). A comparison of the segment of Orf2 that corresponds to module 4 with FAS from rat (18) is illustrated in Fig. 3. It can be seen that, with the exception of a gap of about 400 amino acids in Orf2 and of 600 in FAS, the two sequences align end to end. This indicates an identical order of functional domains in both type I FAS and Orf2 and identifies the putative domains for DH and ER. Functional domains for the KS, AT, KR, and ACP specified by module 4 were identified by multiple alignments of the six modules (13) (Fig. 3). The exact extent of the carboxyl-terminal and amino-terminal domains for DH and ER, respectively, could not be established, but the best conserved regions between module 4 and rat FAS are located at the two ends of the DH + ER domain (Fig. 3), suggesting that the active site domains for DH and ER are located in the amino-terminal and carboxyl-terminal portion of this segment, respectively (13). A similar comparison of the segment of Orf3 corresponding to module 6 with rat FAS (Fig. 3) indicates that the last domain present at the end of Orf3 corresponds to the TE of FAS, as has been previously pointed out (14).

A model for 6dEB synthesis. We propose that each of the *eryA* modules participates in 6dEB formation and is specific for one of the six elongation steps required for production of the full-length chain. The FAS-like functions specified by each module constitute one

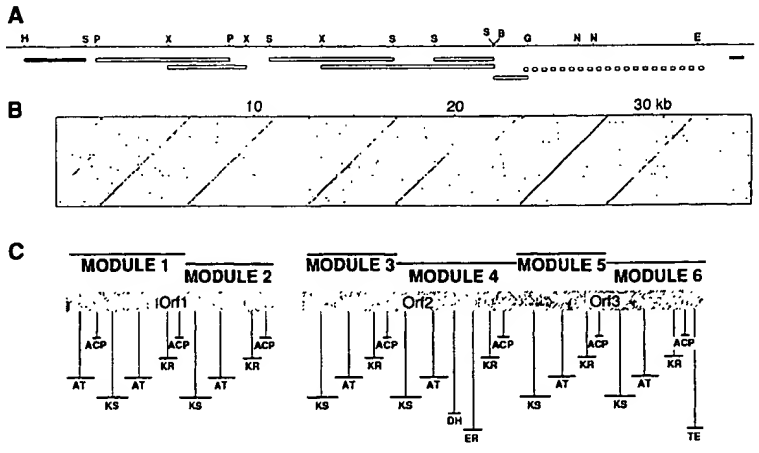

Fig. 2. Organization of *eryA*. (**A**) Gene disruptions. Solid bars refer to EryA$^+$ integrants; open bars to EryA integrants. The dotted bar refers to previously identified *eryA* DNA (7, 8). Only the restriction sites that define the fragments employed for gene disruptions or those mentioned in the text are shown according to the abbreviations B, Bam HI; E, Eco RI; G, Bgl II; H, Hind III; N, Nco I; P, Pvu II; S, Sst I; and X, Xho I. The Eco RI site is 11.5 kb downstream of the Bam HI site of *ermE* (9). DNA fragments were subcloned into pWHM3 (10) by standard methods (33), and integrative transformants of erythromycin-producing *S. erythraea* ER720 (29) were isolated as described (11). The integrated nature of the incoming plasmid and the characterization of the Ery phenotype were performed as described (7). (**B**) Repeated units within *eryA*. Comparison of a 4.5-kb subsegment to the 35-kb segment sequenced with the use of COMPARE-DOTPLOT programs (34) with a window of 50 and a stringency of 32. (**C**) Extension of the six modules, Orf's and putative FAS-like domains. The nucleotide sequence that corresponds to the three polypeptides can be obtained from GenBank under accession numbers M63676 (for Orf1) and M63677 (for Orf's 2 and 3). Putative FAS-like domains within *eryA* ORF's are identified by bars labeled ACP (acyl carrier protein), AT (acyltransferase), DH (dehydratase), ER (enoylreductase), KR (ketoreductase), KS (ketoacyl-ACP synthase), and TE (thioesterase). Bars are placed at their approximate positions; the length reflects the approximate extent of the domains.

functional synthase unit (SU) and the synthesis of 6dEB proceeds by a FAS-like mechanism. At each biosynthetic step, the acyl-KS is formed, the ACP within the same SU is charged with the proper extender unit, and condensation takes place. The resulting β-ketoacyl–ACP undergoes the extent of processing determined by the particular functional domains present within the SU. The processed acyl chain is then transferred from the ACP to the KS of the next SU, which carries all the functions required for the next elongation step. After six cycles, the full-length C_{15} chain has been synthesized through the concerted action of the six SU's, each carrying the specificity for the proper extender unit and for the extent of processing of the β-carbonyl.

The asymmetry of the structure of 6dEB and the differences among the functional domains specified by each module facilitates the assignment of modules to particular steps in the synthesis. Because a methylene group is present at the C-7 position of the 6dEB ring (Fig. 1), corresponding to biosynthesis step 4, this step must be carried out by the SU determined by module 4 (SU4), the only module that specifies DH and ER. Furthermore, the additional AT and ACP domains determined by module 1 are required for formation of propionyl-KS (via propionyl-ACP); thus module 1 is likely to carry out step 1. Similarly, the presence in module 6 of a domain that corresponds to the TE of FAS suggests the participation of this module in transacylation of the full-length C_{15} acyl chain (step 6). Thus, the different components of putative FAS-like activities from the various *eryA* modules suggest that modules 1, 4, and 6 participate in biosynthesis steps 1, 4, and 6, respectively. In addition, although module 3 carries a segment that corresponds to a KR domain, we propose that this KR is nonfunctional, as it lacks the highly conserved VSRRG motif and 9 out of 25 amino acid residues that are found to be invariant in the other five *eryA* KR domains, as well as in the KR domains from chicken (*19*) and rat FAS (Fig. 4). This suggests that SU3 participates in synthesis step 3. Modules 2 and 5 may also participate in the corresponding synthesis steps, so that the genetic order of the six modules corresponds to the order in which the corresponding synthase units are employed in the six elongation steps.

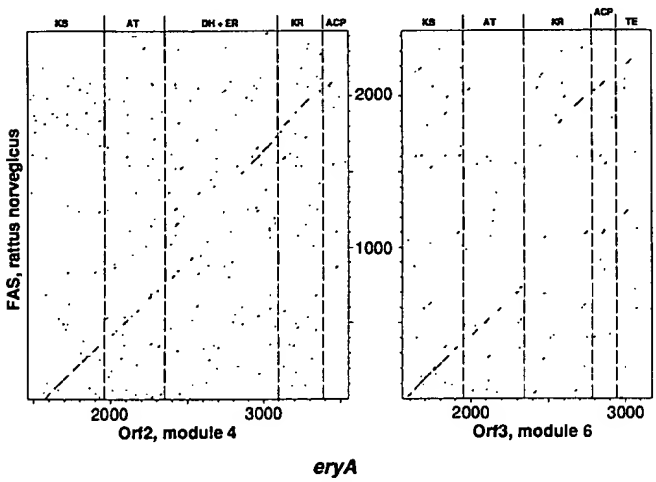

Fig. 3. Comparison of *eryA* with FAS from rat. The translations of the *eryA* segments that correspond to module 4 (left) and module 6 (right) are compared to FAS from rat (*18*) with the use of COMPARE/DOTPLOT programs (*34*) with a window of 30 and a stringency of 16. The proposed extent of each functional domain within *eryA* is marked by vertical dashed lines.

Our model (Fig. 5) displays the proposed intermediates formed after completion of each cycle as enzyme-bound acyl-ACP thioesters. The first ACP from module 1 is acylated by propionate. After condensation of propionate with (2*R*)-mmCoA and β-keto reduction, the first step is completed with formation of the C_5 β-hydroxyacyl-ACP by SU1. The C_5 intermediate is then transferred to the KS of SU2, where (2*S*)-mmCoA is condensed, β-keto reduction takes place, and the C_7 intermediate is transferred from Orf1 to the KS from module 3 in Orf2. Here, (2*R*)-mmCoA is condensed by SU3 and, on the resulting C_9 β-ketothioester, SU4 adds a new (2*R*)-mmCoA unit, with the subsequent full cycle of β-keto reduction, dehydration, and enoylreduction yielding the C_{11} intermediate. The two final elongation steps are carried by SU5 and SU6 in Orf3, both employing (2*S*)-mmCoA units and requiring β-keto reduction, with formation of the C_{13} and then full-length C_{15} chain. Finally, the TE from module 6 transacylates the C_{15} acyl chain from SU6, possibly to CoA, with the resulting acyl-CoA acted upon by a cyclase, which has been proposed to be located 4 kb downstream of module 6 (*20*). Our model is consistent with the results of feeding experiments (*5*) and with the hypothetical macrolactone precursors isolated from some macrolide producers (*6*).

Evidence for the model. We have proposed that each module is involved in a single elongation step, that the choice of extender unit and the extent of processing of the β-carbonyl at each step are determined by the specificity of the corresponding SU and by the functional domains it contains, respectively, and that the linear order of modules corresponds to the succession of elongation steps. Thus, a selective mutation in a given segment of *eryA* should affect a single biochemical event in a predicted elongation step. According to the model, the KR of SU5 is responsible for the introduction of the hydroxyl group appearing at C-5 in the completed polyketide. If such reduction does not take place, and if the succeeding steps in the synthesis, including lactonization, do occur, the compound 5,6-

Fig. 4. Alignments of segments of KR domains from *eryA* Orf's with KR domains from FAS systems. The six *eryA* segments identified by the module that encodes them were manually aligned with a combination of BESTFIT and LINEUP programs (*34*). The resulting consensus was used against the two FAS sequences with PROFILE-SEARCH (*34*). Invariant residues are marked by dashes and gaps by dots. The best conserved segments between *eryA* Orf's and FAS's are enclosed by boxes. The shaded box (top) refers to the proposed NADPH (*16*) binding sites. Bars below alignment of KR's indicate major deviations from possible consensus in the KR from module 3.

```
chicken   KSYIITGGLGGFGLELAQWLI.ERGAQKLVLTSRSGIRTGYQAKCVREWKALGIQVLVSTSDVGTLEGTQLLIEEALKLG.PVGGIFNLAVVL
rat       KSYII---LGGF-LEL-RW-V.LR-AQRLV-TS-S-IRTGYQAKHVR-WRRQ-IHVLVSTSNVSSLEGARALIAEATKLG.PVGGVFNL-MVL
module 6  GTALV---TGAL-GHV-RH-A.RC-VEDLV-VS-R-VDAPGAAELEA-LVAL-AKTTITACDVADREQLSKLLEELRGQGRPVRTVVHT-GVP
module 5  GTVLV---TGGI-AHV-RW-A.RS-AEHLV-LG-R-ADAPGASELRE-LTAL-TGVTIAACDVADRARLEAVLAAERAEGRTVSAVMHA-GVS
module 4  GTVLI---TGTL-RLL-RH-VTEH-VRHLL-VS-R-ADAPGSDELRA-IEDL-ASAEIAACDTADRDALSALLDG.LPR..PLTGVVHA-GVL
module 2  GTILV---TAGL-AEV-RW-A.GR-AEHLA-VS-R-PDTEGVGDLTA-LTRL-ARVSVHACDVSSREPVRELVHGLIEQGDVVRGVVHA-GLP
module 1  GTVLV---TGGV-GQI-RW-A.RR-APHLL-VS-S-PDADGAGELVA-LEAL-ARTTVAACDVTDRESVRELLGG.IGDDVPLSAVFHA-ATL
module 3  GTVLV--AASPV-DQLVRW-A.DR-AERLV-.....AGACPGDDLLAAVEEA-A.SAVVAQDAA...ALREALG.....DEPVTALVHAGTLT

chicken   KDAMIENQTPELFWEVNKPKYSGTLHLDWVTRKKCPDLD..YFVVFSSVSCGRGNAGQSNYGFANSAMERICEQRHHDGLPGLAVQWGAIGDV
rat       RDAMLENQTPELFQDVNKP-YNGTLN-DRATREACPELD..Y-VA---VSCGRGNAGQSN-GFA-STMERICEQ-RHD-LPGLAVQ--AIGDV
module 6  ESRPLHEIGE..LESVCAA-VTGARL-D....ELCPDAE..T-VL---GAGVWGSANLGA-SAA-AYLDALAHR-RAE-RAATSVA--AWAGE
module 5  TSTPLDDLTEAEFTEIADV-VRGTVN-D....ELCPDLD..A-VL---NAGVWGSPGLAS-AAA-AFLDGFARR-RSE-APVTSIA--LWAGQ
module 4  ADGLVTSIDEPAVEQVLRA-VDAAWN-H....ELTANTGLSF-VL---AASVLAGPGQGV-AAA-ESLNALAAL-RTR-LPAKALG--LWAQA
module 2  QQVAINDMDEAAFDEVVAA-AGGAVH-D....ELCSDAE..L-LL---GAGVWGSARQGA-AAG-AFLDAFARH-RGR-LPATSVA--LWAAG
module 1  DDGTVDTLTGERIERASRA-VLGARN-H....ELTRELDLTA-VL---FASAFGAPGLGG-APG-AYLDGLAQQ-RSD-LPATAVA--TWAGS
module 3  NFGSISEVAPEEFAETIAA-TALLAV-D....EVLGDRAVEREVYC--VAGIWGGAGMAA-AAGSAYLDALAEHHRAR-RSCTSVA-TPWALP
```

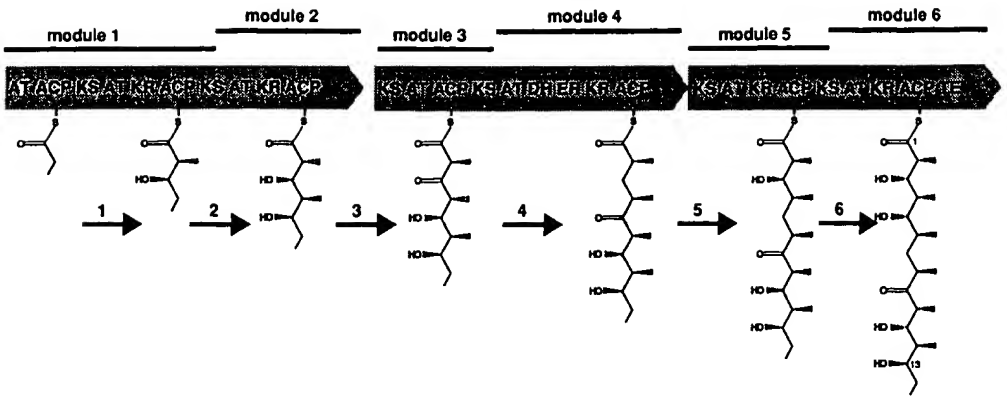

Fig. 5. Model for 6-deoxyerythronolide B formation. The starter unit and the intermediates formed at the end of each elongation cycle are shown as the corresponding ACP thioesters. The FAS-like functions responsible for each elongation step are shown inside each Orf, with the corresponding modules above the three Orf's. Numbered arrows refer to the six elongation steps. Lactonization of the C_{15} chain between carbons 1 and 13 results in the formation of 6-deoxyerythronolide B (Fig. 1).

dideoxy-5-oxoerythronolide B should be produced. An 813-bp in-frame deletion was introduced in vitro into the DNA segment of module 5 that specifies the KR domain, and the mutant allele was introduced into *S. erythraea* to replace the wild-type counterpart in the chromosome (*21*). The resulting strain AKR5 produced 5,6-dideoxy-3α-mycarosyl-5-oxoerythronolide B as the major component, as well as minor amounts of 5,6-dideoxy-5-oxoerythronolide B (Fig. 6). This indicates that the KR of SU5 is responsible for β-keto reduction in synthesis step 5 and provides direct evidence for the model. It also indicates that hydroxylation at C-6, which normally takes place on 6dEB, does not occur on the corresponding 5-deoxy-5-oxo derivative and confirms the substrate flexibility of the enzyme machinery which does not require hydroxyl groups at C-6 as recently reported (*20*) or at C-5 and C-6, as our data show. Because desosamine attachment occurs at the C-5-hydroxyl group of the lactone ring, 5,6-dideoxy-3α-mycarosyl-5-oxoerythronolide B could not be further processed, and in this way the compound was accumulated by strain AKR5.

Minor quantities of the compounds 2-norerythromycins have been isolated from the *S. erythraea eryA* strain 9EI41 when transformed with a cosmid library of heterologous DNA. This family of compounds resulted from the incorporation of malonate instead of methylmalonate at the sixth condensation step (*22*). Strain 9EI41 carries an 87-bp in-frame deletion in the segment of module 6 that specifies the AT domain (7), thus this result is consistent with the involvement of the AT from module 6 in synthesis step 6, as predicted by the model, and suggests that the DNA that complemented the *eryA* mutation in strain 9EI41 encoded a malonyl-CoA transferase. The yield of 5-deoxy-5-oxoerythronolide B is about 100-fold greater than that of 2-norerythromycin and thus highlights the requirement for precise interaction among the various domains in the multifunctional polypeptides. The observation of alternate compounds illustrates the lack of absolute specificity of the KS of SU6, which condensed a malonate into the growing chain in this case and employed a β-keto-C_{13} substrate in the case of strain AKR5.

Implications of the model. The proposed model explains how the synthesis of an asymmetric polyketide such as 6dEB is pro-

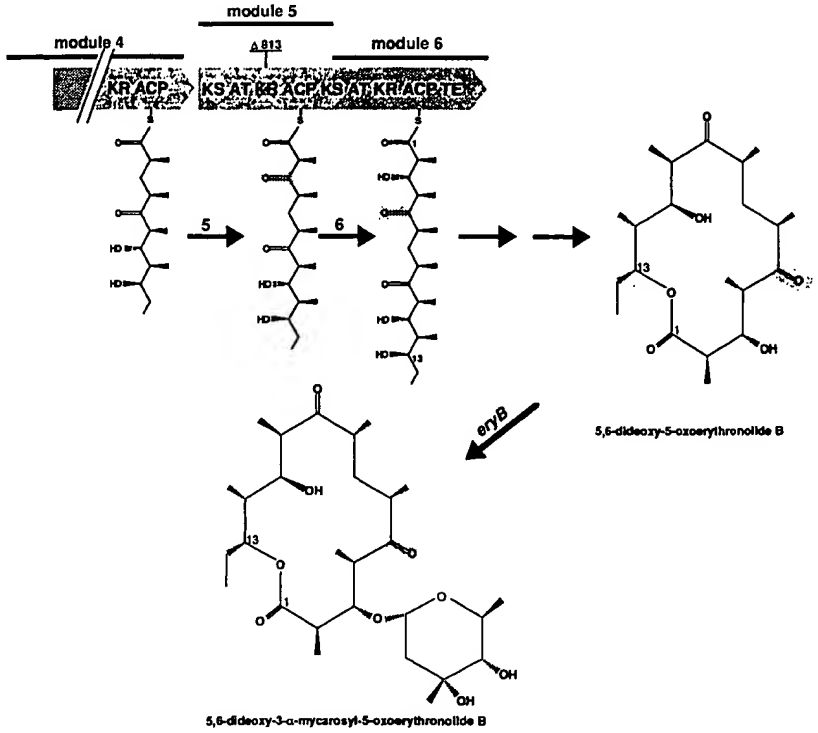

Fig. 6. Pathway for formation of 5,6-dideoxy-3α-mycarosyl-5-oxoerythronolide B. The altered Orf3 that carried an inactive KR in SU5 and the corresponding 813-bp deletion in strain AKR5 are shown on top. The keto group that remained after elongation step 5 is highlighted throughout the synthesis. Strain AKR5 was fermented for 3 days according to conditions described (*22*) and the broth was extracted three times with ethyl acetate. The dried extract was digested in toluene and chromatographed on a column of silica gel (Merck Kiesel gel 60, 70-230 mesh) and eluted with a step gradient of isopropanol in toluene. The 5 and 7 percent isopropanol eluates were combined and concentrated to an oily residue (800 mg), which was digested in a solvent system consisting of *n*-heptane:chloroform:ethanol (10:10:1, by volume) and chromatographed on a Sephadex LH-20 column that has been packed and developed with the same solvent system. Fractions (9 to 13) were combined, concentrated, and crystallized from a mixture of *n*-heptane and ethyl acetate to give 5,6-dideoxy-3α-mycarosyl-5-oxoerythronolide B (45.2 mg, melting point 163° to 164°C). Structure was deduced by spectral analysis and confirmed by single crystal x-ray to diffraction analysis. Fractions (15 to 17) were combined and crystallized from ethyl acetate to yield 5,6-dideoxy-5-oxoerythronolide B (5.3 mg) with spectral characteristics identical to those of an authentic sample (*32*).

grammed in the producing microorganism. Rather than through the utilization of a hypothetical enzyme capable of altering its specificity in response to variations in length and side groups of the growing acyl chain, the correct polyketide is made through the concerted action of six SU's, each devoted to a single FAS-like cycle. The final molecular structure is programmed by (i) use of a single and distinct module encoding the information required for each elongation step and (ii) selection of the correct SU to be transacylated by the growing chain.

The current model for palmitic acid formation in type I FAS systems calls for transacylation of the growing chain after each cycle between two distinct but identical polypeptides (*2*). Six different SU's participate in 6dEB synthesis, however. Some mechanism must thus ensure that the correct SU is transacylated after each cycle by the required substrate. The fidelity of this transacylation may depend on appropriate steric juxtaposition of ACP and KS domains of adjacent SU's or on the substrate specificity of the KS. The observation that a β-hydroxyl C_{13} chain is transferred from the ACP of SU5 to the KS of SU6 in wild-type Orf3, whereas the same transfer occurs with a β-keto C_{13} chain by AKR5 Orf3, suggests that the precise intramolecular transfer is independent of the acyl chain substrate structure. The close proximity of ACP and KS domains that are part of two adjacent SU's in the primary structure of each polypeptide may provide the necessary steric juxtaposition required for correct intramolecular transacylation. Intermolecular transfer could also result from having a single KS domain sterically available to each ACP domain if the three polypeptides assemble in a specific manner in a hypothetical PKS complex. It is possible that such precise geometric arrangement of the three polypeptides is facilitated by the colinearity of genetic and biochemical order in 6dEB formation.

The presence of an AT specific for propionyl-CoA in SU1 and the formation of 2-norerythromycin in a strain defective in the AT of SU6 suggest a direct function for the AT in the choice of the correct extender unit. Furthermore, the apparent lack of specificity of the KS and ACP domains of SU6 for the extender unit employed, (2*S*)-mmCoA in the wild-type and malonylCoA in the 2-norerythromycin–producing derivative of strain 9EI41, suggests that these domains do not function in extender unit choice.

Homopolymeric aromatic polyketides made through successive condensations of the same extender unit are synthesized by type II PKS's in bacteria (*17*) and by a type I enzyme in fungi (*23*). Synthesis of these simple polyketides requires the participation of no more than six polypeptides in the type II systems (*3*). According to our model, the fidelity of 6dEB synthesis is ensured by the involvement of one SU for each elongation step. It is difficult to conceive how the synthesis of this complex polyketide, which requires the proper interactions among six different SU's for a total of 28 functional FAS-like domains, could be achieved in a type II system. The occurrence in prokaryotes of type I or type II systems may thus relate to the complexity of the polyketide made. The preliminary evidence suggesting that *S. erythraea* FAS is a type II enzyme (*24*) is consistent with this proposal and indicates that type I and type II systems may coexist within one organism. The proposed model for the synthesis of 6dEB and the organization of the corresponding PKS in large multifunctional polypeptides are analogous to the nonribosomal template synthesis of peptide antibiotics and to the organization of the relative synthetases (*25*).

Genes for the formation of the macrolides tylosin (*26*), spiramycin (*27*), and avermectin (*28*) also contain repeated units. We propose, therefore, that the genes for the biosynthesis of macrolides and possibly of other complex polyketides, for example polyenes and polyethers, are also organized in modules in a manner similar to *eryA*. If, as illustrated by the example of the KR from module 5 in *eryA*, each SU exhibits some substrate flexibility, an alteration of a functional domain of the PKS may result in the formation of a novel polyketide of predicted structure. Modular organization of PKS genes and colinearity of the genetic and biochemical order permits the opportunity to make select genetic alterations in polyketide-producing organisms and thus offers possibilities for the tailored production of polyketide structures.

REFERENCES AND NOTES

1. A. J. Birch and F. W. Donavan, *Aust. J. Chem.* 6, 360 (1953).
2. S. J. Wakil, *Biochemistry* **28**, 4523 (1989).
3. D. A. Hopwood and D. H. Sherman, *Annu. Rev. Genet.* **24**, 37 (1990).
4. S. Omura and Y. Tanaka, in *Macrolide Antibiotics: Chemistry, Biology and Practice*, S. Omura, Ed. (Academic Press, New York, 1984), p. 199.
5. D. E. Cane and C. Yang, *J. Am. Chem. Soc.* **109**, 1255 (1987); S. Yue, J. S. Duncan, Y. Yamamoto, C. R. Hutchinson, *ibid.*, p. 1253.
6. K. Kinoshita, S. Takenaka, M. Hayashi, *J. Chem. Soc. Chem. Commun.* **1988**, 943 (1988); M. L. B. Huber *et al.*, *Antimicrob. Agents Chemother.* **34**, 1535 (1990).
7. J. S. Tuan *et al.*, *Gene* **90**, 21 (1990).
8. J. M. Weber *et al.*, *J. Bacteriol.* **172**, 2372 (1990).
9. M. J. Bibb, G. R. Janssen, J. M. Ward, *Gene* **38**, 215 (1985); H. Uchiyama and B. Weisblum, *ibid.* **38**, 103 (1985).
10. J. Vara, M. Lewandowska-Skarbek, Y.-G. Wang, S. Donadio, C. R. Hutchinson, *J. Bacteriol.* 171, 5872 (1989).
11. J. M. Weber and R. Losick, *Gene* **68**, 173 (1988).
12. Plasmid DNA was transformed into *S. erythraea* ER720 (*29*) and two integrants from each transformation were analyzed for the production of bioactive compounds and for the presence of ethyl acetate–extractable metabolites by thin-layer chromatography (TLC) as described (*7*). The EryA phenotype was characterized by the lack of known erythromycin intermediates and by the absence of detectable bioactivity in the ethyl acetate extract (*7*).
13. M. J. Staver, S. Donadio, J. S. Tuan, L. Katz, in preparation.
14. J. Cortes, S. H. Haydock, G. A. Roberts, D. J. Bevitt, P. F. Leadlay, *Nature* **348**, 176 (1990).
15. For a recent compilation of active site motifs for these activities, see (*3*) and (*14*). Abbreviations for the amino acid residues are A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
16. N. S. Scrutton, A. Berry, R. N. Perham, *Nature* **343**, 38 (1990).
17. M. J. Bibb, S. Biro, H. Motamedi, J. F. Collins, C. R. Hutchinson, *EMBO J.* **8**, 2727 (1989); D. H. Sherman *et al.*, *ibid.*, p. 2717.
18. M. Schweizer, K. Takabayashi, T. Laux, K.-F. Beck, R. Schreglmann, *Nucleic Acids Res.* **17**, 567 (1989).
19. Z. Yuan, W. Liu, G. G. Hammes, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6328 (1988); K. P. Holzer, W. Liu, G. G. Hammes, *ibid.* **86**, 4387 (1989); S. S. Chirala *et al.*, *J. Biol. Chem.* **264**, 3750 (1989).
20. J. M. Weber, J. O. Leung, S. J. Swanson, K. B. Idler, J. B. McAlpine, *Science* **252**, 114 (1991).
21. The 9.5-kb Bam HI–Xho I segment between coordinates 22.0 and 31.5 (Fig. 2A) was subcloned into the Bam HI–Sal I fragment of pUC19 (*30*) to generate pAIBX9. The internal 813-bp Nco I fragment of this plasmid was removed and in the resulting construct the pUC19 portion was replaced with pWHM3 (*10*) by Eco RI–Hind III digestion. The resulting plasmid pAKR5.2 was introduced into *S. erythraea* ER720 protoplasts and integrative transformants were isolated (*11*). Thiostrepton-sensitive colonies (five) were isolated from one integrative transformant as described (*31*) and three were found upon genomic Southern (DNA) blot hybridization to have retained the deleted *eryA* segment in the chromosome. These colonies, which upon TLC analysis (*7*) showed no visible erythromycin intermediates but produced a product that corresponded to authentic 5,6-dideoxy-5-oxoerythronolide B (*32*), were named strain AKR5.
22. J. B. McAlpine *et al.*, *J. Antibiot. (Tokyo)* **40**, 1115 (1987); in preparation.
23. J. Beck, S. Ripka, A. Seigner, E. Schlitz, E. Schweizer, *Eur. J. Biochem.* **192**, 487 (1990).
24. R. S. Hale, K. N. Jordan, P. F. Leadlay, *FEBS Lett.* **224**, 133 (1987).
25. H. Kleinkauf and H. von Dohren, *Annu. Rev. Microbiol.* **41**, 259 (1987); *Eur. J. Biochem.* **192**, 1 (1990).
26. R. H. Baltz and E. T. Seno, *Annu. Rev. Microbiol.* **42**, 547 (1988).
27. M. A. Richardson *et al.*, *J. Bacteriol.* **172**, 3790 (1990).
28. D. J. MacNeil, K. M. Gewain, C. L. Ruby, T. MacNeil, P. H. Gibbons, *J. Cell. Biochem.* **14A**, (suppl.) 92, (1990).
29. J. DeWitt, *J. Bacteriol.* **164**, 969 (1985).
30. J. Vieira and J. Messing, *Gene* **19**, 259 (1982).
31. S. Donadio and C. R. Hutchinson, *Gene*, in press.
32. J. R. Martin and T. J. Perun, *Biochemistry* **7**, 1728 (1968).
33. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
34. J. Devereux, P. Haeberli, O. Smithies, *Nucleic Acids Res.* **12**, 387 (1984).

35. We thank D. Brown for help with fermentations; D. Whittern and P. Hill for spectral determinations; S. Spanton for x-ray diffraction analysis; and C. R. Hutchinson, R. Losick, and P. Youngman for critical reading of the manuscript.

25 January 1991; accepted 20 March 1991